Semi-Annual Status Report

submitted to
NASA
Goddard Space Flight Center
Greenbelt, Maryland

Error Control Techniques for Satellite and Space Communications
Grant Number NAG5-557
June 1, 1985 – May 31, 1988

Principal Investigator:
Daniel J. Costello, Jr.
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556

January, 1988

Summary of Progress

During the period June 1, 1987 – November 30, 1987, progress was made in the following areas:

1) Concatenated Codes Using Bandwidth Efficient Trellis Inner Codes

A paper summarizing our work on the performance of bandwidth efficient trellis inner codes using two-dimensional MPSK signal constellations in a NASA concatenated coding system has been accepted for publication in the IEEE Transactions on Communications [1]. A preliminary version of this paper was included in our previous report (June 1987). The final version of the paper was recently re-submitted. Reprints will be sent when they become available.

A second paper summarizing the performance of multi-dimensional trellis coded MPSK modulation as inner codes in a concatenated coding system has also been submitted for publication to the IEEE Transactions on Communications [2]. Multi-dimensional trellis codes are ideally suited for concatenated coding systems since we can match the number of information bits per branch on the inner code to the symbol size of the outer code. Thus each incorrectly decoded branch in the inner code results in a single symbol error for the outer code. This property of multi-dimensional trellis codes is exactly analogous to the property of unit-memory convolutional codes, first discovered by Lee [3], which made them natural choices for use in a concatenated system. Our results indicate that the symbol-oriented nature of multi-dimensional trellis codes can provide an improvement of up to 1dB in the overall performance of a concatenated coding system when these codes replace bit-oriented two-dimensional trellis inner codes of similar decoding complexity. A preprint of our paper is included as Appendix A of this report.

Our work on concatenated coding using bandwidth efficient trellis codes was conducted by Dr. Robert H. Deng, a former post-doctoral research associate supported by the Grant.

2) Bounds on the Minimum Free Euclidean Distance of Bandwidth Efficient Trellis Codes

A paper summarizing our work on lower bounds on the minimum free Euclidean distance of bandwidth efficient trellis codes has been accepted for publication in the IEEE Transactions on Information Theory [4]. A preliminary version of this paper was included in our previous report (June 1987). The final version of the paper was recently re-submitted. Reprints will be sent when they become available.

A weakness of the bound presented in the above paper is that it is not tight at short constraint lengths, the most useful range for practical codes. We are currently working on a new bound which is much tighter at short constraint lengths. This bound can be used to predict the achievable free distance for practical constraint lengths and signal constellations. Code searches can then be conducted using the most promising signal constellations. A summary of our work on this new bound is included as Appendix B of this report [5]. A

full length paper is being prepared for submission to the IEEE Transactions on Information Theory.

Our work on minimum distance bounds for bandwidth efficient trellis codes is being conducted by Mr. Marc Rouanne, a Ph.D. student supported by the Grant.

3) Performance Analysis of Bandwidth Efficient Trellis Codes on Channels with Phase Jitter

Bandwidth efficient trellis coded phase modulation (TCPM) is a powerful means of achieving high data rates on bandlimited satellite channels. The performance of these coding/modulation schemes is usually analyzed assuming an ideal AWGN channel model. Many real channels, however, also suffer from phase jitter due to synchronization inaccuracies in the phase recovery loop. This phase jitter results in a time-varying phase error by which the signal constellation is rotated. We have begun an examination of the effect of this phase error on the performance of TCPM schemes. A summary of our initial results is included as Appendix C of this report [6].

We have found that codes which are optimum for the ideal AWGN channel may perform quite poorly on a chennel subject to phase jitter. We are able to measure the performance degradation of these codes in decibels as a function of the ratio of the phase noise power of the receiver to the thermal noise power of the channel. For example, Ungerboeck's [7] 64-state, rate 2/3, 8-PSK scheme, which achieves a 4.77 dB coding gain over uncoded QPSK on an ideal channel, suffers a 1.25 dB degradation in performance when the phase noise power is 20% of the thermal noise power. We are conducting a computer search for codes which are optimum on channels with phase jitter. The performance of these codes will then be compared to codes which were designed for the ideal channel. The results of this code search will be included in our next report.

Our work on bandwidth efficient trellis codes on phase jitter channels is being conducted by Mr. Christian Schlegel, a Ph.D. student supported by the Grant.

4) Construction of Bandwidth Efficient Trellis Codes

We have continued our work on the construction of multi-dimensional bandwidth efficient trellis codes with MPSK modulation. A paper submitted to the IEEE Transactions on Information Theory was included in our February 1987 semi-annual progress report. This paper has now been accepted for publication and is currently being revised for re-submission [8]. Upon initial submission, the paper attracted quite a bit of attention from the Information Theory community. Two other researchers (G. Ungerboeck from the IBM communications research laboratory in Zurich, Switzerland and S. S. Pietrobon from the South Australian Institute of Technology, Adelaide, South Australia) contacted us about combining our results with similar results which they had obtained. The revised paper now contains contributions from three research groups and will make an outstanding contribution to the field of coded modulation. In addition to the original constructions of multi-dimensional trellis codes for 8-PSK modulation, new multi-dimensional codes have been added for 16-PSK modulation and multi-

dimensional 8- and 16-PSK codes have been constructed which are invariant (transparent) to rotations of the signal set when used with differential encoding. This latter property allows continued reliable communication when the receiver tracking loop locks onto the wrong signal phase. Among the advantages of multi-dimensional codes over conventional two-dimensional constructions are (i) the ability to achieve higher data rates, (ii) better coding gains, (iii) the ability to achieve transparency to discrete phase rotations of the signal set, and (iv) better performance as inner codes in a concatenated system. Reprints of the final version of this paper will be submitted when they become available.

Our work on the construction of multi-dimensional trellis codes was initiated by Mr. Alain LaFanechere, a former student supported by the Grant, and by Dr. Robert H. Deng and is now being continued, with emphasis on the construction of codes which are invariant to phase rotations of the signal set, by Mr. Steven S. Pietrobon, a new Ph.D. student supported by the Grant.

In addition to the construction of multi-dimensional trellis codes, we have also begun work on some hueristic construction algorithms for bandwidth efficient trellis codes with larger constraint lengths. These constructions are sub-optimum, but are designed to produce good long bandwidth efficient codes which are suitable for use with sequential decoding. The large constraint lengths make an exhaustive search for optimum codes impractical, but we are encouraged that good long codes can be constructed using our algorithms. A summary of our initial results is included as Appendix D of this report [9]. The work is being conducted by Mr. Marc Rouanne.

5) Parity Retransmission Hybrid ARQ Using Convolutional Codes

A paper summarizing our work using convolutional codes with parity retransmission hybrid ARQ has been accepted for publication in the IEEE Transactions on Communications [10]. The final version of the paper is currently in preparation. A preprint is included as Appendix E of this report. This paper complements work done by Lin and Yu with block codes [11]. The paper includes an unique analysis of the performance of these codes on bursty channels. A major result is that throughput efficiency is shown to improve as the channel becomes more bursty in nature. This is consistent with the design philosophy of type-II hybrid ARQ schemes, which are intended for use on bursty channels. The paper also demonstrates that the convolutional code scheme achieves a very high reliability. This work was initiated by Dr. Laurent R. Lugand, a former Ph.D. student supported by the Grant, and was completed by Dr. Robert H. Deng.

The current paper discusses only the use of rate 1/2 convolutional coding schemes. Higher throughputs can be achieved by using rate 2/3 and 3/4 schemes, although the retransmission protocol becomes more complex. We will discuss the performance of these higher rate schemes in a future report.

References

[1] R. H. Deng and D. J. Costello, Jr., "High Rate Concatenated Coding Systems Using Bandwidth Efficient Trellis Inner Codes," *IEEE Trans. Commun.*, to appear.

[2] R. H. Deng and D. J. Costello, Jr., "High Rate Concatenated Coding Systems Using Multi-Dimensional Bandwidth Efficient Trellis Inner Codes," *IEEE Trans. Commun.*, submitted for publication, July 1987.

[3] L. N. Lee, "Concatenated Coding Systems Employing a Unit-Memory Convolutional Code and a Byte-Oriented Decoding Algorithm," *IEEE Trans. Commun.*, COM-25, pp. 1064-1074, October 1977.

[4] M. Rouanne and D. J. Costello, Jr., "A Gilbert Lower Bound on the Minimum Free Euclidean Distance of Trellis Coded Modulation," *IEEE Trans. Inform. Th.*, to appear.

[5] M. Rouanne and D. J. Costello, Jr., "An Expurgated Lower Bound on the Free Euclidean Distance of Trellis Codes," Proceedings Allerton Conference on Communication, Control, and Computing, Monticello, IL, October 1987.

[6] C. Schlegel and D. J. Costello, Jr., "Bandwidth Efficient Coding on Channels with Phase Jitter," submitted to the 1988 IEEE International Symposium on Information Theory, November 1987.

[7] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," *IEEE Trans. Inform. Theory*, IT-28, pp. 55-67, January 1982.

[8] R. H. Deng, S. S. Pietrobon, G. Ungerboeck, A. LaFanechere, and D. J. Costello, Jr., "Multi-Dimensional Trellis Coded Phase Modulation," *IEEE Trans. Inform. Th.*, to appear.

[9] M. Rouanne and D. J. Costello, Jr., "Construction of Good Trellis Codes," submitted to the 1988 IEEE International Symposium on Information Theory, November 1987.

[10] L. R. Lugand, D. J. Costello, Jr., and R. H. Deng, "Parity Retransmission Hybrid ARQ Using Rate 1/2 Convolutional Codes on a Non-Stationary Channel," *IEEE Trans. Commun.*, to appear.

[11] S. Lin and P. S. Yu, "A Hybrid ARQ Scheme with Parity Retransmission for Error Control of Satellite Channels," *IEEE Trans. Commun.*, COM-30, pp. 1701-1719, July 1982.

Appendix A

High Rate Concatenated Coding Systems Using Multi-Dimensional Bandwidth Efficient Trellis Inner Codes

High Rate Concatenated Coding Systems Using Multi-Dimensional Bandwidth Efficient Trellis Inner Codes*

Robert H. Deng and Daniel J. Costello, Jr.
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

July 10, 1987

Abstract

In a previous paper [1], a concatenated coding system employing 2-dimensional (2-D) trellis coded MPSK inner codes and Reed-Solomon outer codes for application in high speed satellite communication systems was proposed. This paper extends the results of [1] to systems using symbol-oriented multi-dimensional trellis coded MPSK inner codes.

The concatenated coding systems will be divided into two classes according to their achievable effective information rates. The first class uses multi-dimensional trellis coded 8PSK inner codes and achieves effective information rates around 1 bit/Dimension (spectral efficiency 2 bps/Hz). The second class employs multi-dimensional trellis coded 16PSK inner codes and provides effective information rates around 1.5 bits/Dimension (spectral efficiency 3 bps/Hz). Both classes provide significant coding gains over an uncoded reference system with the same effective information rate as the coded system.

*This work was supported by NASA Grant NAG5-557.

I. Introduction

In [1], a concatenated coding system with 2-dimensional (2-D) trellis coded MPSK (TCMPSK) inner codes and Reed-Solomon (RS) outer codes for application in high speed satellite communication systems was proposed. It was argued there that TCMPSK inner codes along with soft decision Viterbi decoding plays two important roles in a concatenated coding system:

1. Compensating for the bandwidth expansion introduced by the outer code;

2. Compressing the random errors on the inner channel into symbol errors which can be corrected by a symbol-error-correcting outer code, such as an RS code.

In trellis (convolutional) inner code/RS outer code concatenated coding systems, such as those in [1] and [2] which employ a soft decision Viterbi decoder for the inner code, it is unlikely that the beginning of a decoding error burst is aligned with the boundary between two RS symbols. This fact was first observed by Lee [3] for binary convolutional inner code/RS outer code concatenated coding systems and led to the discovery of symbol-oriented unit memory inner convolutional codes. This observation leads us to consider using symbol-oriented multi-dimensional (multi-D) TCMPSK inner codes rather than bit-oriented 2-D TCMPSK inner codes. A typical concatenated coding system is shown in Figure 1. The outer code is an (N, K) RS code with $N = 2^b - 1$ and symbols over $GF(2^b)$. The inner code is rate $R_1 = b/n$, $2^\nu - state$, multi-D TCMPSK, where b, the number of

information bits entering the inner encoder per encoding interval, is chosen to equal the RS code symbol size.

Encoding is performed in two stages. An information sequence of Kb bits is divided into K symbols of b bits each, and each b-bit symbol is regarded as an element of $GF(2^b)$. These K symbols are used as inputs to the RS encoder. The output of this encoder is an N-symbol codeword which is symbol-interleaved and then serially encoded by the trellis encoder with b input bits per encoding interval. Decoding is accomplished in the reverse order. The inner channel is assumed to be an additive white Gaussian noise (AWGN) channel with single sided power spectrum N_0. The inner code is decoded by a Viterbi decoder without demodulator output quantization. The outer decoder is an errors-only RS decoder.

The concatenated coding systems will be divided into two classes according to their achievable effective information rate. Class 1 systems use multi-D TC8PSK inner codes and achieve effective information rates around 1 bit/Dimension (spectral efficiency 2 bps/Hz). Class 2 systems employ multi-D TC16PSK inner codes and achieve effective information rates around 1.5 bits/Dimension (spectral efficiency 3 bps/Hz). Their performance is studied in sections II and III, respectively.

II. Systems Employing Multi-D TC8PSK Inner Codes

In this section we study the performance of concatenated coding systems with the multi-D TC8PSK schemes constructed in [4] as inner codes. For any positive integer $L \geq 2$, a rate $R_1 = b/3L, 2L - D$ TC8PSK scheme has

an effective information rate in information bits per dimension (bits/D)

$$R_{eff}^{(1)} = b/2L \geq 1 \; bit/D, \quad 2L \leq b \leq 3L-1. \tag{1}$$

Then the overall effective information rate of the concatenated coding system

$$R_{eff} = R_{eff}^{(1)} \cdot \frac{K}{N} < R_{eff}^{(1)} \tag{2}$$

is around 1 bit/D.

Let P_s denote the symbol error probability into the outer decoder. Assuming bounded distance decoding of the outer code, the decoded bit error rate (BER) at the output of the outer decoder is closely approximated by

$$P_b \approx \frac{d}{2N} \sum_{i=t+1}^{N} \binom{N}{i} P_s^i (1-P_s)^{N-i}, \tag{3}$$

where d is the minimum Hamming distance of the RS code and $t = \lfloor (d-1)/2 \rfloor$ is its symbol-error-correcting capability [5].

The performance of the concatenated coding system will be measured in terms of the coding gain at a given target BER P_b over an uncoded reference system with the same effective information rate R_{eff} as the coded system, i.e., $2^{2R_{eff}}$ PSK modulation.[1] (This comparison was suggested by Forney [6].) The BER for the reference system is then given by

$$P_b \approx \frac{1}{R_{eff}} Q\left(\sqrt{\frac{\Delta_0^2 R_{eff} E_b}{N_0}}\right), \tag{4}$$

[1]Since $2^{2R_{eff}}$ may not be an integer, this uncoded reference system may be only hypothetical. It is used for comparison purposes since it has exactly the same effective information rate as the coded system.

where E_b/N_0 is the channel "information bit energy-to-noise power density ratio", and

$$\Delta_0^2 = 2 - 2\cos\left(2\pi/2^{2R_{eff}}\right) \tag{5}$$

is the normalized minimum squared Euclidean distance of the $2^{2R_{eff}}$ PSK signal set.

The coding gain of a concatenated coding system over the uncoded reference system, denoted by γ, is found as follows. We first find the $(E_b/N_0)_{uncoded}$ required to achieve a target BER from (4). Then we find the $(E_b/N_0)_{coded}$ required to achieve the same BER using (3), where the symbol error probability P_s is obtained by computer simulation of the inner decoder. The coding gain is then given by

$$\gamma = 10\log_{10}\frac{\left(\frac{E_b}{N_0}\right)_{uncoded}}{\left(\frac{E_b}{N_0}\right)_{coded}} dB. \tag{6}$$

Fig. 2(a) shows the coding gain at $P_b = 10^{-6}$ and $P_b = 10^{-9}$ with respect to R_{eff} for a concatenated coding system with $R_1 = 5/6$, 4-D TC8PSK inner codes and $N = 2^5 - 1 = 31$ RS outer codes. Results are given for inner codes with 4 and 8 states. The gain (loss) of the reference system over QPSK is also plotted in the figure. The coding gain over QPSK of a concatenated coding system can be found by adding γ to the gain (loss) over QPSK of the reference system with the same R_{eff}. To achieve higher effective information rates and larger coding gains, more powerful RS outer codes can be employed. Fig. 2(b) shows the coding gain with respect to R_{eff} for a concatenated coding system with the same inner codes as in Fig. 2(a) but with $N = 2^{10} - 1 = 1023$ RS outer codes. The RS symbol size is

twice the number of input bits to the inner encoder per encoding interval in this case.

Fig. 3(a) depicts the coding gain for a concatenated coding system with $R_1 = 7/9$, 6-D TC8PSK inner codes and $N = 2^7 - 1 = 127$ RS outer codes. Higher effective information rates can be achieved by a concatenated coding system employing $R_1 = 8/9$, 6-D TC8PSK inner codes and $N = 2^8 - 1 = 255$ RS outer codes. Its coding gain is shown in Fig. 3(b). Both systems achieve roughly the same coding gains with a 4-state inner code.

Fig. 4 shows the coding gain for a concatenated coding system with $R_1 = 8/12$, 8-D TC8PSK inner codes and $N = 2^8 - 1 = 255$ RS outer codes. The three inner codes have the same minimum squared Euclidean distance d_f^2, but decreasing path multiplicity as the number of trellis states increases. Note that only about 0.1 dB more gain is obtained with every doubling of the number of trellis states.

III. Systems Employing Multi-D TC16PSK Inner Codes

The overall effective information rate of a concatenated coding system with multi-D TC8PSK inner codes is always less than 1.5 bits/D. To achieve higher rates, multi-D TC16PSK must be used as inner codes. For any positive integer $L \geq 2$, the $R_1 = b/4L, 2L - D$ TC16PSK schemes constructed in [7] have an effective information rate

$$R_{eff}^{(1)} = b/2L \geq 1.5 \; bits/D, \;\; 3L \leq b \leq 4L - 1. \tag{7}$$

From (2), the overall effective information rate of the concatenated coding system is around 1.5 bits/D.

Due to the symbol-oriented nature of the inner codes, the concatenated coding system performance can be estimated using formula calculations as well as by simulation. Let $S(d_f^2)$ denote the total number of symbol errors associated with paths that are distance d_f from the correct path, normalized by the number of decoding intervals on each path. Then $S(d_f^2)$ is upper bounded by the number of paths that are distance d_f from the correct path, $P(d_f^2)$, a parameter known for most TCM codes. To see this, let $P_i(d_f^2)$ be the number of paths of length i decoding intervals that are distance d_f from the correct path, and let n_i be the maximum number of symbol errors associated with these paths. Then it follows that

$$S(d_f^2) \leq \sum_i \left(\frac{P_i(d_f^2) \cdot n_i}{i} \right) \leq \sum_i \left(\frac{P_i(d_f^2) \cdot i}{i} \right) = \sum_i P_i(d_f^2) = P(d_f^2), \tag{8}$$

since the number of symbol errors along any path of length i is bounded by n_i and n_i is bounded by i. The symbol error probability P_s to the outer

decoder, for large values of E_b/N_0, can be approximated by

$$P_s \approx S(d_f^2)Q\left(\sqrt{\frac{d_f^2 R_{eff} E_b}{N_0}}\right) \leq P(d_f^2)Q\left(\sqrt{\frac{d_f^2 R_{eff} E_b}{N_0}}\right), \qquad (9)$$

where R_{eff} is the overall effective information rate of the concatenated coding system. The final decoded BER P_b can be found by using (9) in (3).

Since a close approximation to P_s using (9) requires a high E_b/N_0 ratio, or equivalently $P_s << 1$, in the following we only consider RS outer codes with d = 3 and 5, since these are sufficient to achieve decoded BER's in the range $10^{-6} \sim 10^{-9}$. Fig. 5 compares the performance obtained by the formula calculations to that obtained through computer simulations for a concatenated coding system with an $R_1 = 5/6$, 4-state, 4-D TC8PSK inner code and an $N = 31$ RS outer code. It is seen that the formula calculations and the simulations are very close at $P_b \leq 2 \times 10^{-5}$ for d = 3 and $P_b \leq 2 \times 10^{-7}$ for $d = 5$.[2]

The coding gains obtained by formula calculations vs. the inner code constraint length ν are shown in Figs. 6(a) and 6(b) for systems with 4-D TC16PSK inner codes and in Figs. 7(a) and 7(b) for systems with 6-D TC16PSK inner codes. The coding gains of the inner codes alone are also shown in the figures for comparison.

The advantage of concatenated coding over the inner code alone is obvious. With a d = 3 RS outer code, the concatenated coding system offers $0.75 \sim 1.25\ dB$ more coding gain at $P_b = 10^{-6}$ and $1.25 \sim 1.75\ dB$ more

[2]The formula calculations are expected to be more accurate for larger RS code lengths N, say $N = 127$, for then the RS code rate is higher and lower values of P_s are needed to achieve $P_b = 10^{-6}(P_s \approx 10^{-4})$ and $P_b = 10^{-9}(P_s \approx 10^{-5})$.

coding gain at $P_b = 10^{-9}$, respectively, than the inner code alone (3 dB more coding gain asymptotically). With a d = 5 RS outer code, the concatenated coding system offers $1.25 \sim 2\ dB$ more coding gain at $P_b = 10^{-6}$ and $2.25 \sim 2.5\ dB$ more coding gain at $P_b = 10^{-9}$, respectively, than the inner code alone (4.77 dB more coding gain asymptotically).

IV. Conclusions

We have studied the performance of concatenated coding systems with symbol-oriented multi-D TCMPSK inner codes. The advantages of using symbol-oriented multi-D TCMPSK inner codes are best seen by comparing the coding performance with that of concatenated coding systems employing bit-oriented inner codes. Fig. 8 shows the performance of two concatenated coding systems. System 1 uses an Ungerboeck $R_1 = 2/3$ ($R_{eff}^{(1)} = 1$ bit/D), 16-state, 2-D TC8PSK [8] inner code and a (255,223) RS outer code. System 2 employs an $R_1 = 8/12$ ($R_{eff}^{(1)} = 1$ bit/D), 4-state, 8-D TC8PSK [4] inner code and the same outer code. Both systems have an effective information rate $R_{eff} = 0.875$ $bits/D$. Ungerboeck's 16-state code has a 4.13 dB asymptotic coding gain, while the 4-state, 8-D code has only a 3 dB asymptotic coding gain and a much larger path multiplicity. However, system 2 is inferior to system 1 by only $0.05 \sim 0.11$ dB at $P_b = 10^{-6} \sim 10^{-9}$. Thus, the symbol-oriented nature of the 4-state, 8-D TC8PSK inner code provides an improvement of more than 1 dB in overall performance. Moreover, the 4-state code is simpler to decode than the 16-state code.

To further justify this observation, Fig. 9 shows another system performance comparison. In Fig. 9, system 1 uses an $R_1 = 8/9$ ($R_{eff}^{(1)} = 1.33$ $bits/D$), 4-state, 2-D periodic time-varying trellis coded (PTVTC) 8PSK [9] inner code and a (255, 201) RS outer code. System 2 uses an $R_1 = 8/9$, 4-state, 6-D TC8PSK [4] inner code and the same outer code. Both systems have $R_{eff} = 1.05$ $bits/D$. Both inner codes have a 2.9 dB

asymptotic coding gain and roughly the same path multiplicity. However, system 2 offers 0.6 dB more coding gain than system 1, which is entirely due to the symbol-oriented nature of the $R_1 = 8/9$, 4-state, 6-D TC8PSK inner code.

From the performance studies presented above, we can draw a number of conclusions.

1. The symbol-oriented nature of multi-D TCMPSK inner codes can provide an improvement of up to 1 dB in the overall performance of a concatenated coding system when these codes replace bit-oriented 2-D TCMPSK inner codes.

2. Most of the coding gain can be obtained by using $4 \sim 16 - state$ inner codes. Therefore, choosing inner codes with a small number of trellis states increases the data transmission rate (by reducing the number of decoder computations) with only a slight sacrifice in system performance.

3. The path multiplicity of the inner code exerts less influence on the performance of a concatenated coding system than on the performance of the inner code by itself (see Fig. 7(a) at $\nu = 4$ and Fig. 7(b) at $\nu = 2$). This can be explained as follows. Using Forney's [10] rule of thumb, the path multiplicity degrades the performance of a trellis code at $P_b = 10^{-5}$ by 0.25 dB for every increase in the multiplicity by a factor of 2. However, since the errors in the output of a Viterbi decoder are highly bursty, the information bit errors along a path are

concentrated into only a few symbol errors. After deinterleaving, the symbol errors will be corrected with high probability by the RS outer code. This fact provides a basis for statement 2 in section I.

References

[1] R. H. Deng and D. J. Costello, Jr., "High rate concatenated coding systems using bandwidth efficient trellis inner codes," submitted to *IEEE Trans. Commun.*, 1987.

[2] J. P. Odenwalder, "Optimal decoding of convolutional codes," Ph.D. Dissertation, School of Engineering and Applied Sciences, University of California, Los Angeles, 1970.

[3] L.-N. Lee, "Concatenated coding systems employing a unit-memory convolutional code and a byte-oriented decoding algorithm," *IEEE Trans. Commun.*, Vol. COM-25, pp. 1064-1074, Oct. 1977.

[4] A. LaFanchere, R. H. Deng, and D. J. Costello, Jr., "Multidimensional trellis coded phase modulation using unit-memory and partial-unit-memory convolutional codes," submitted to *IEEE Trans. Inform. Theory*, 1987.

[5] S. Lin and D. J. Costello, Jr., *Error-Control Coding: Fundamentals and Applications*, Prentice-Hall, Englewood Cliffs, N.J., 1983.

[6] G. D. Forney, Jr., private communication, 1987.

[7] R. H. Deng and D. J. Costello, Jr., "Multi-D trellis coded 16 PSK modulation," submitted to *IEEE Trans. Inform. Theory*, 1987.

[8] G. Ungerboeck, "Channel coding with multilevel/phase signals,"

IEEE Trans. Inform. Theory, Vol. IT-28, pp. 56-67, January 1982.

[9] F. Hemmati and R. J. F. Fang, "Low complexity coding methods for high data rate channels," Comsat laboratories technical note, Feb. 1984.

[10] G. D. Forney, Jr., "Coset codes I," submitted to *IEEE Trans. Inform. Theory*, 1986.

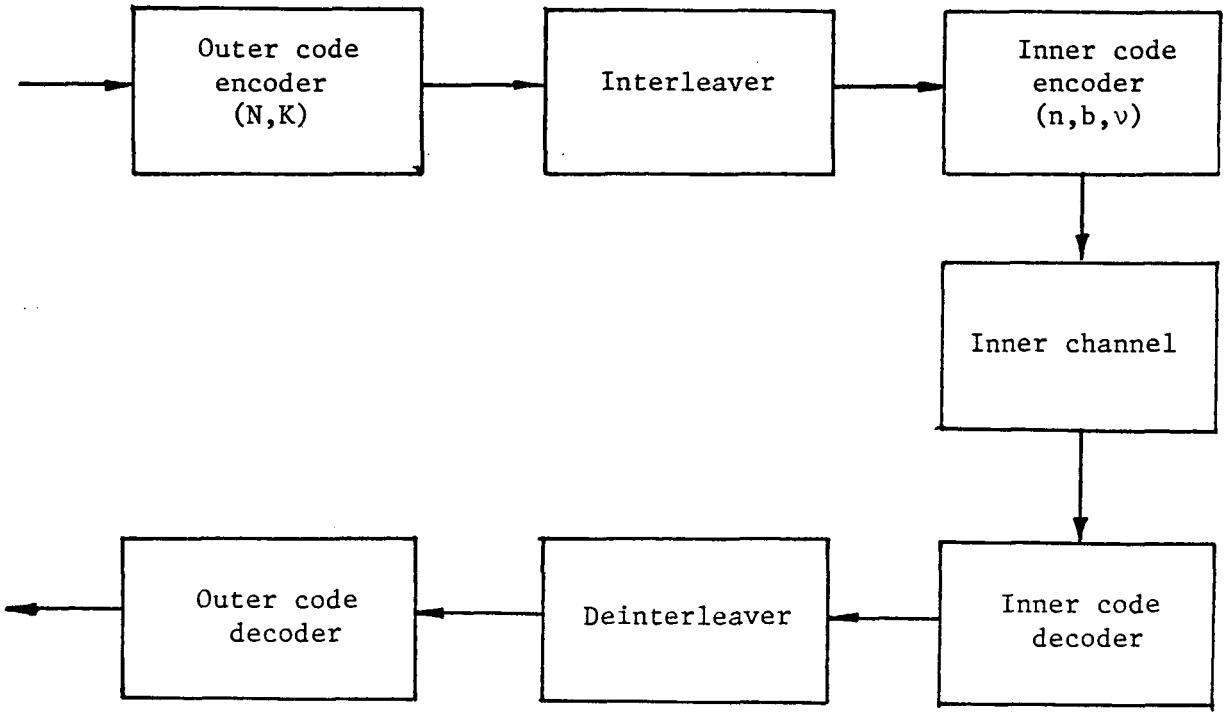


Fig. 1 A concatenated coding system.

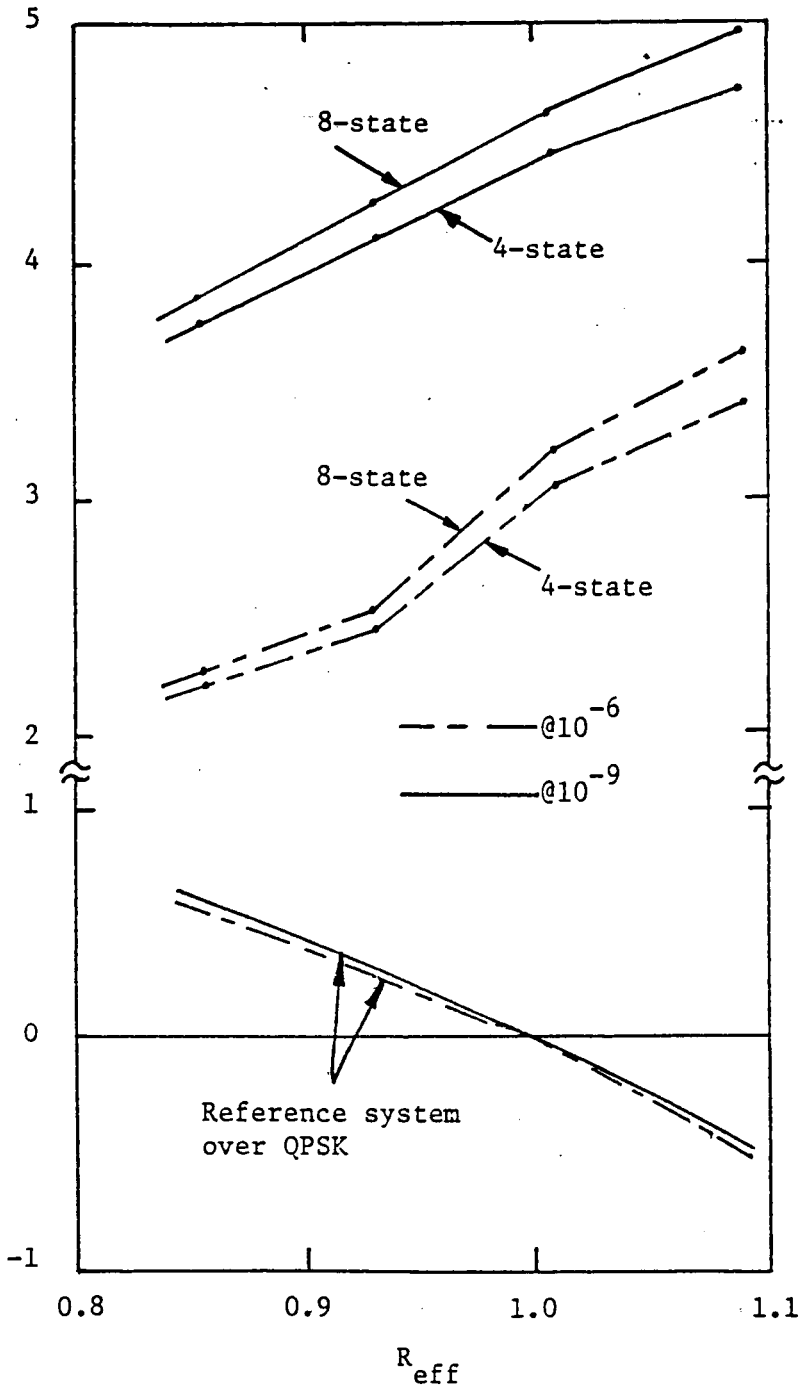


Fig. 2(a) Coding gain vs. R_{eff} for a concatenated coding system with an $R_1 = 5/6$, 4-D TC8PSK inner code and an N = 31 RS outer code with d = 5 ∿ 11.

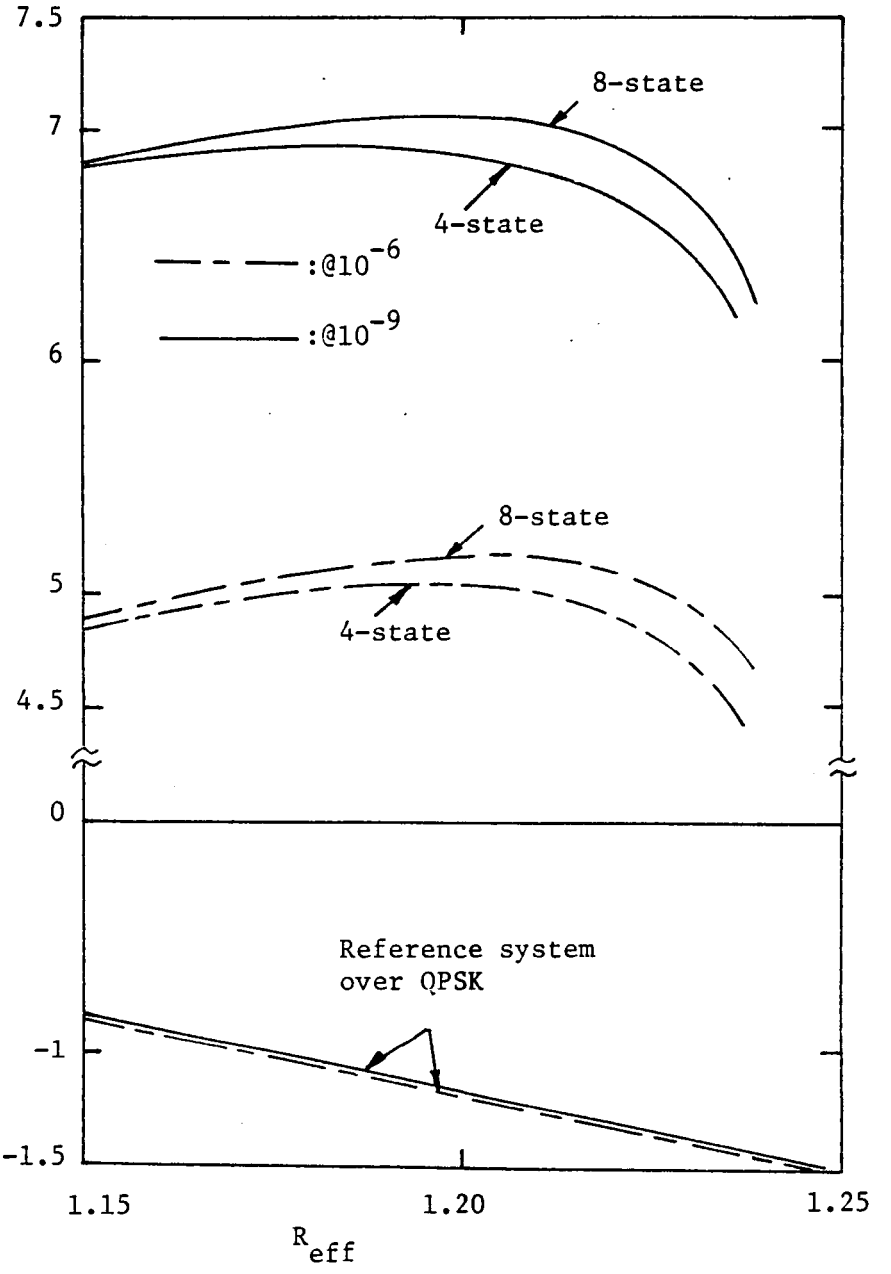


Fig. 2(b) Coding gain vs. R_{eff} for a concatenated coding system with an R_1 = 5/6, 4-D TC8PSK inner code and an N = 1023 RS outer code with d = 13 ∿ 81.

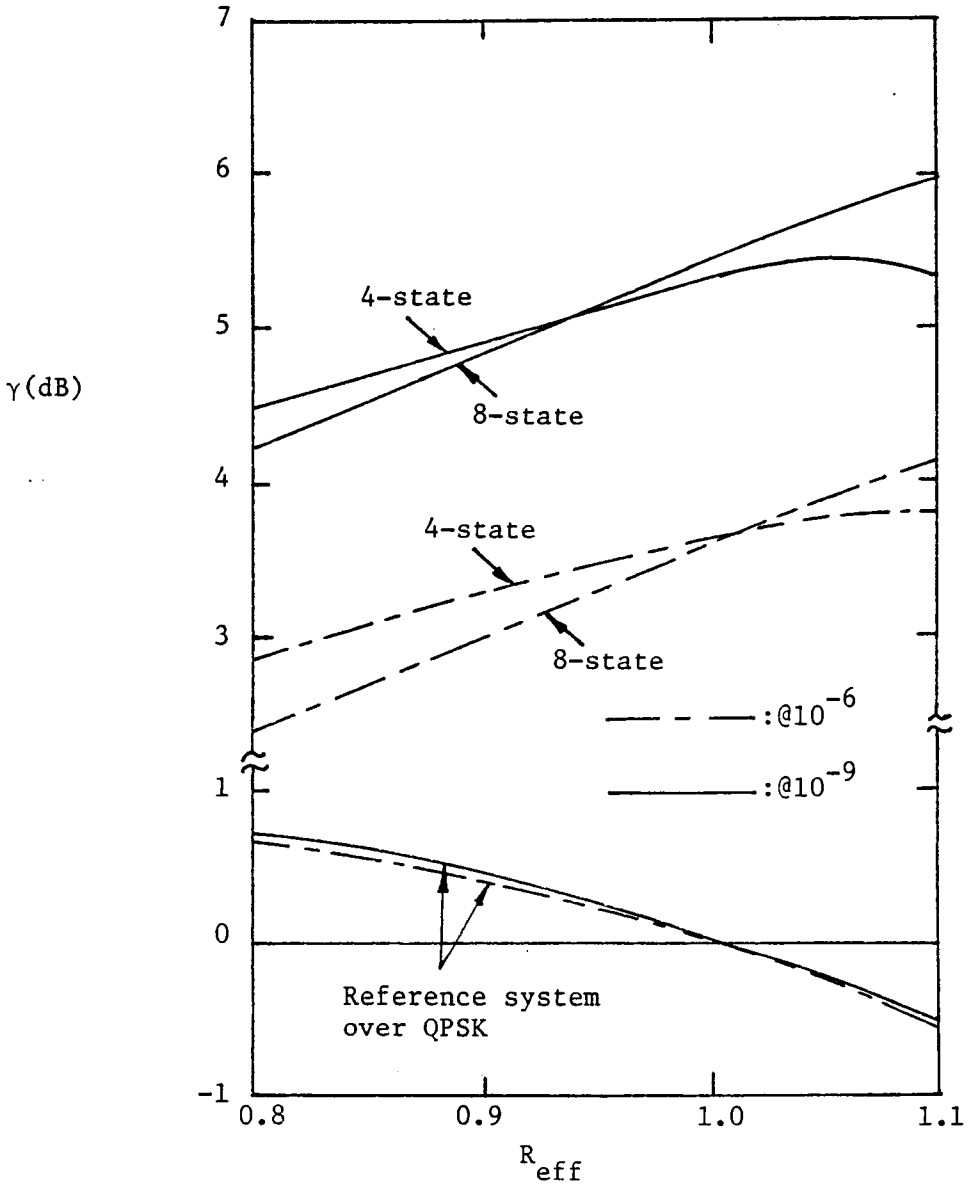


Fig. 3(a) Coding gain vs. R_{eff} for a concatenated coding system with an $R_1 = 7/9$, 6-D TC8PSK inner code and an N = 127 RS outer code with d = 7 ∿ 41.

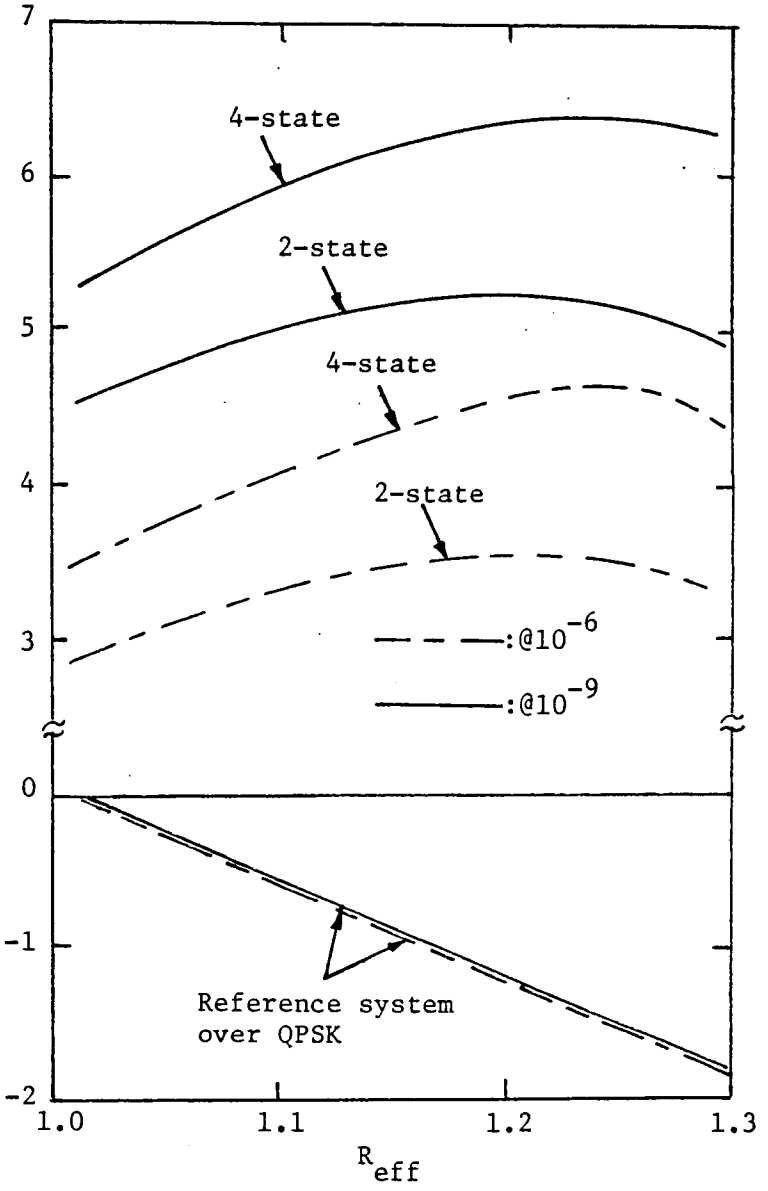


Fig. 3(b) Coding gain vs. R_{eff} for a concatenated coding system with an R_1 = 8/9, 6-D TC8PSK inner code and an N = 255 RS outer code with d = 11 ∿ 61.

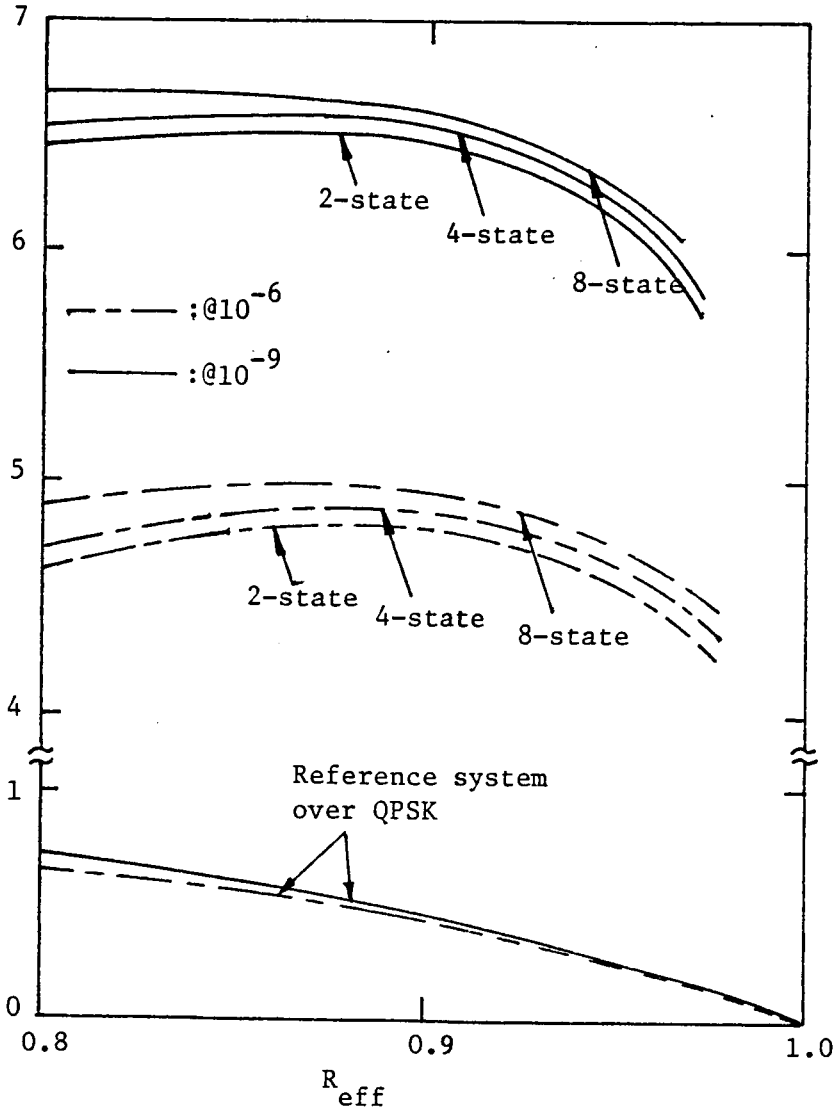


Fig. 4 Coding gain vs. R_{eff} for a concatenated coding system with an R_1 = 8/12, 8-D TC8PSK inner code and an N = 255 RS outer code with d = 11 ∿ 51.

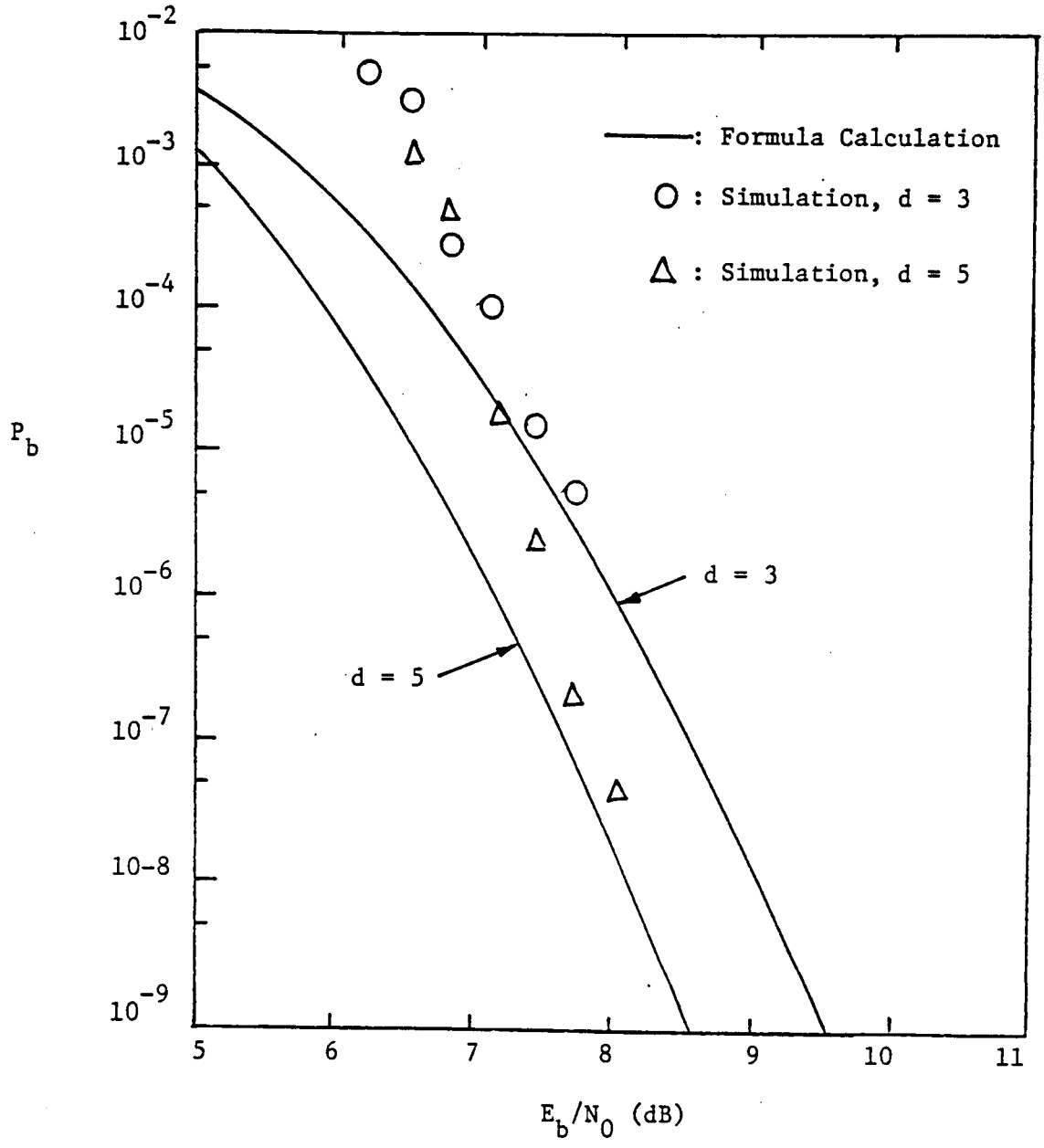


Fig. 5 Performance of a concatenated coding system with an R_1 = 5/6, 4-state, 4-D TC8PSK inner code and an N = 31 RS outer code.

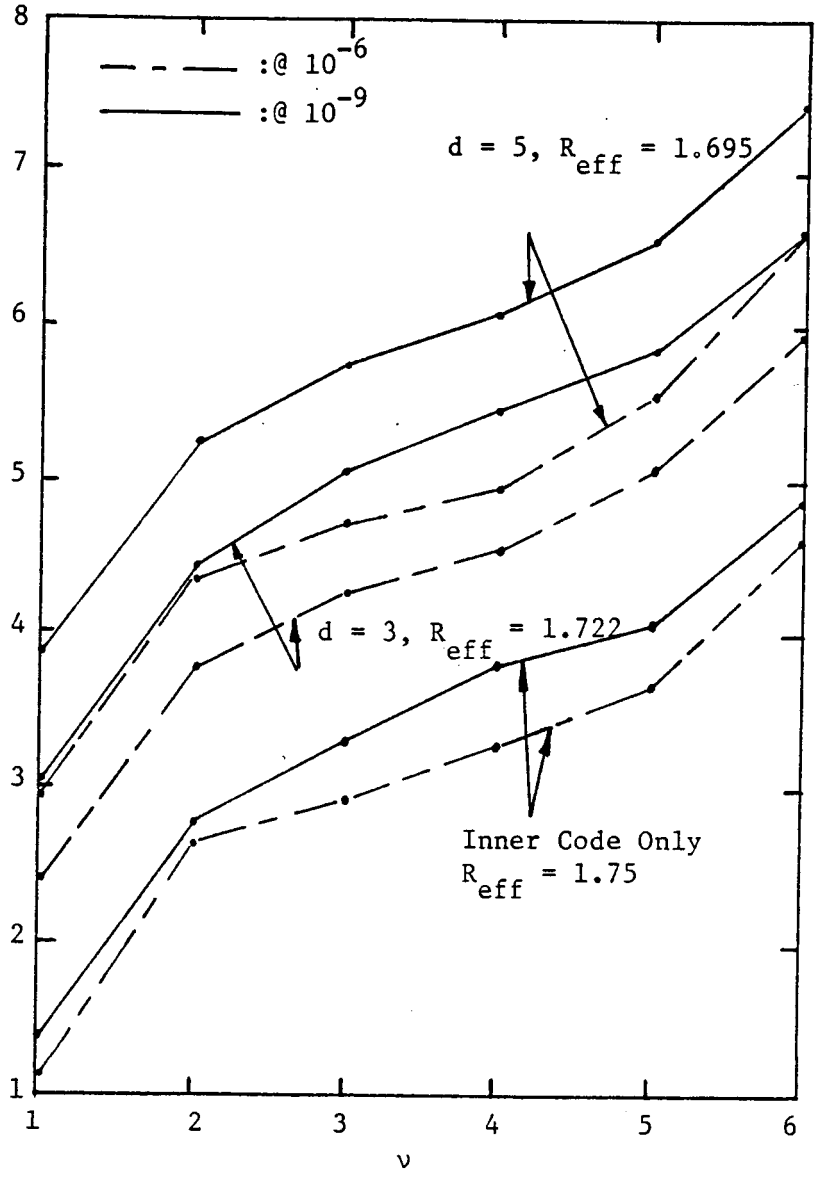


Fig. 6(a) Coding gain vs. ν for a concatenated coding system with an R_1 = 7/8, 4-D TC16PSK inner code and an N = 127 RS outer code.

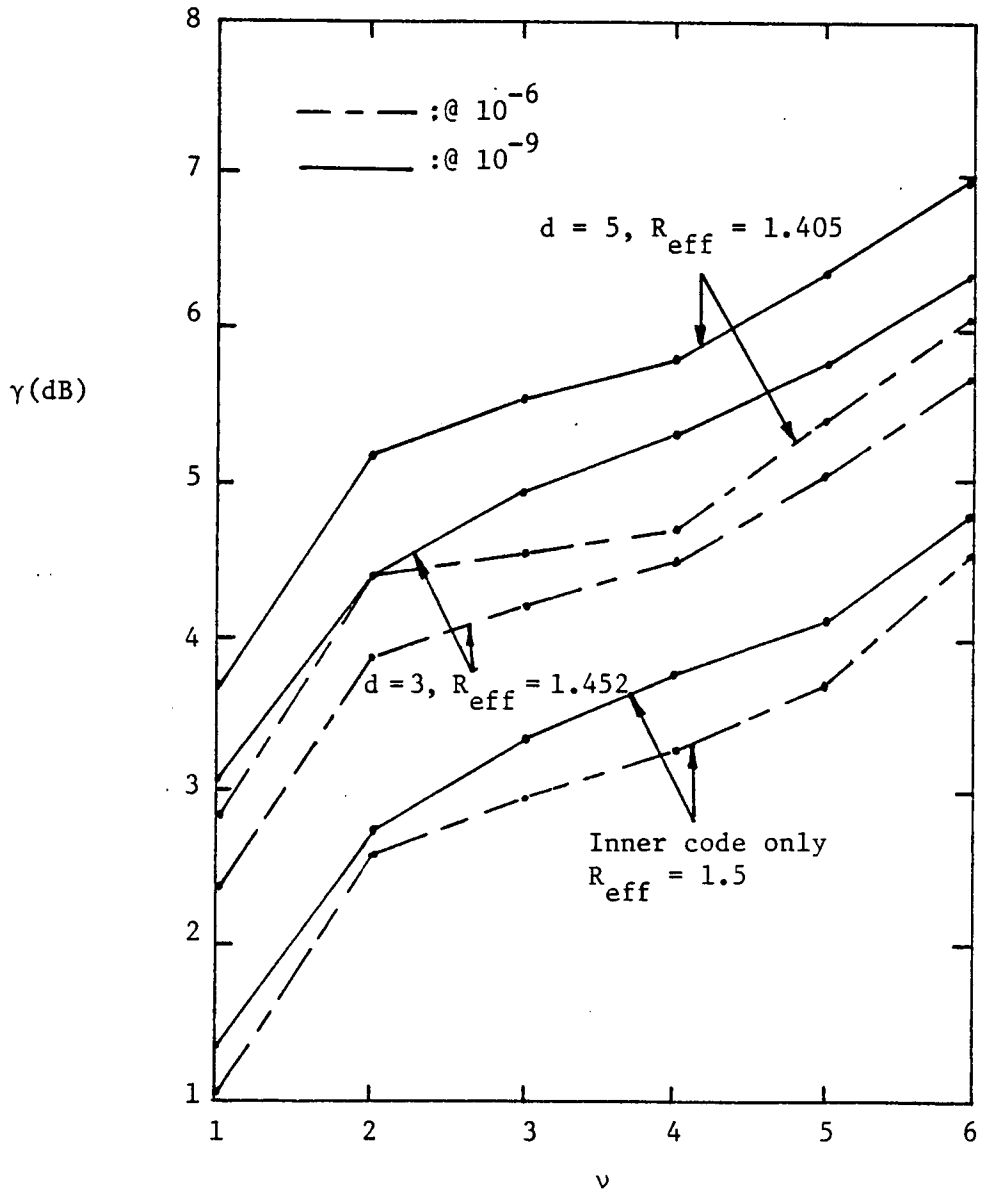


Fig. 6(b) Coding gain vs. ν for a concatenated coding system with an R_1 = 6/8, 4-D TC16PSK inner code and an N = 63 RS outer code.

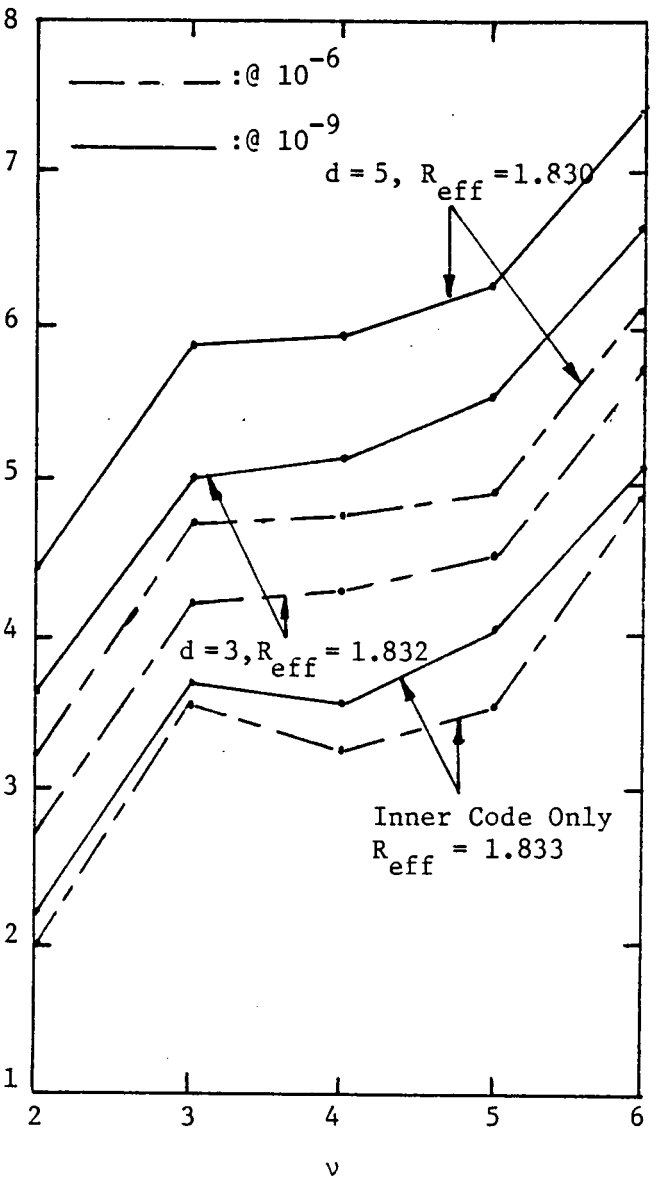


Fig. 7(a) Coding gain vs. ν for a concatenated coding system with an R_1 = 11/12, 6-D TC16PSK inner code and an N = 2047 RS outer code.

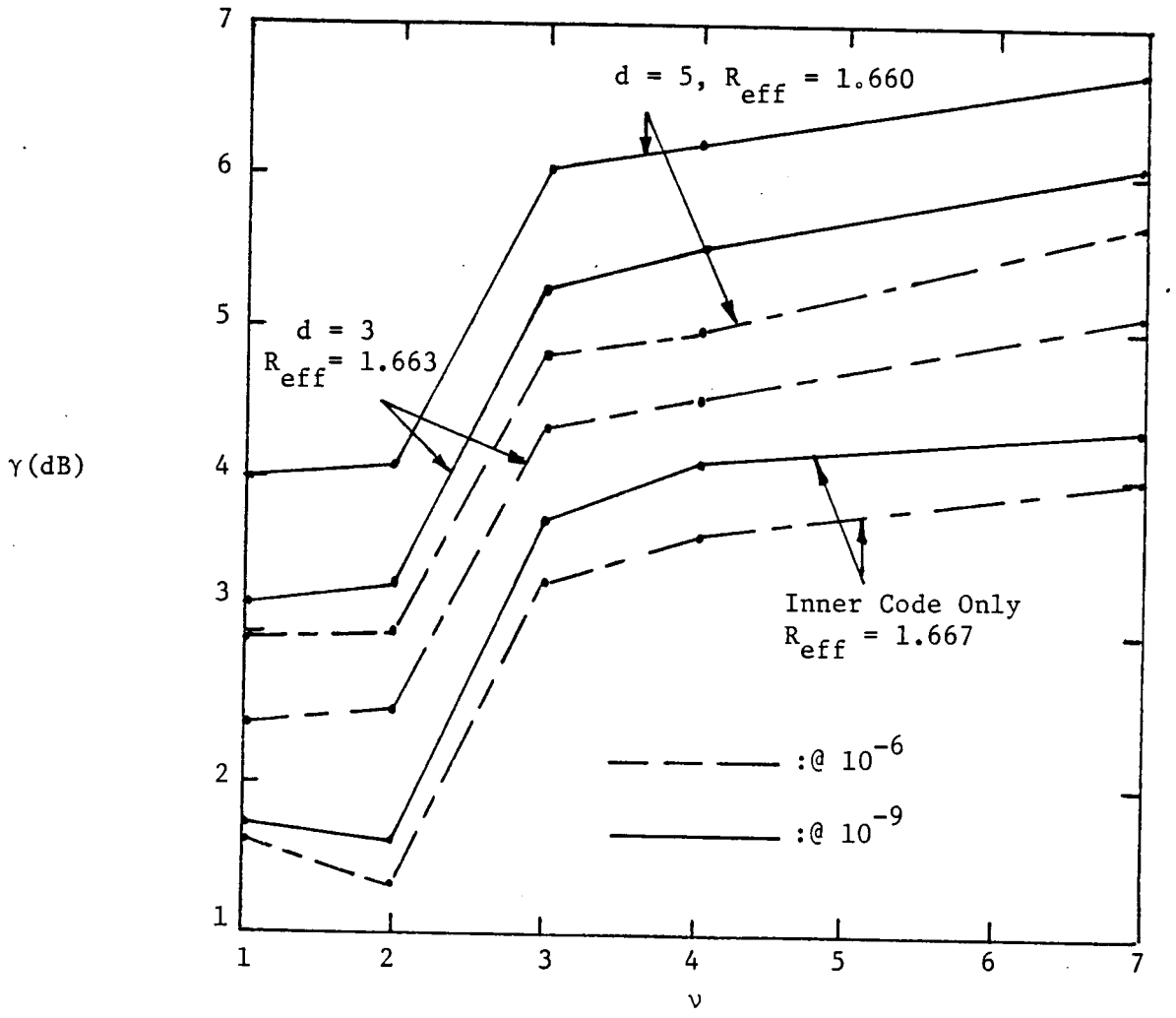


Fig. 7(b) Coding gain vs. ν for a concatenated coding system with an $R_1 = 10/12$, 6-D TC16PSK inner code and an N = 1023 RS outer code.

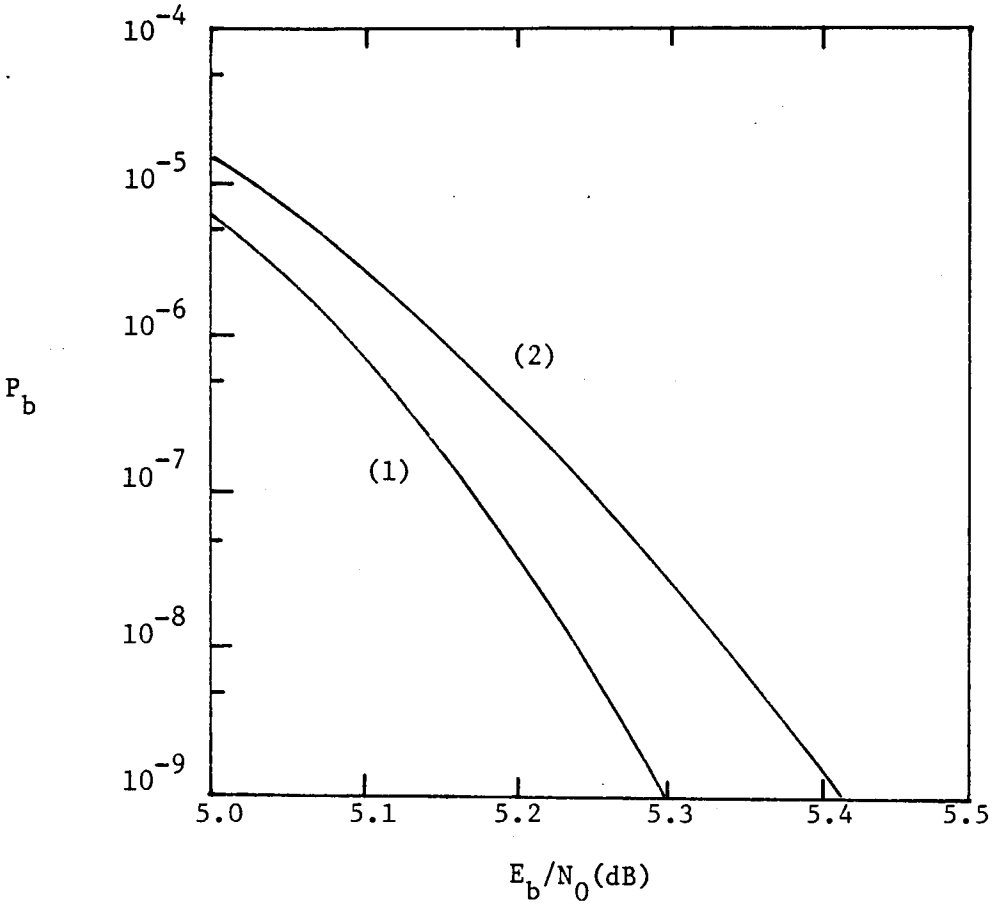


Fig. 8 Performance comparison for concatenated coding systems with

(1) An R_1 = 2/3, 16-state, 2-D TC8PSK inner code and an (255,223) RS outer code;

(2) An R_1 = 8/12, 4-state, 8-D TC8PSK inner code and the same outer code.

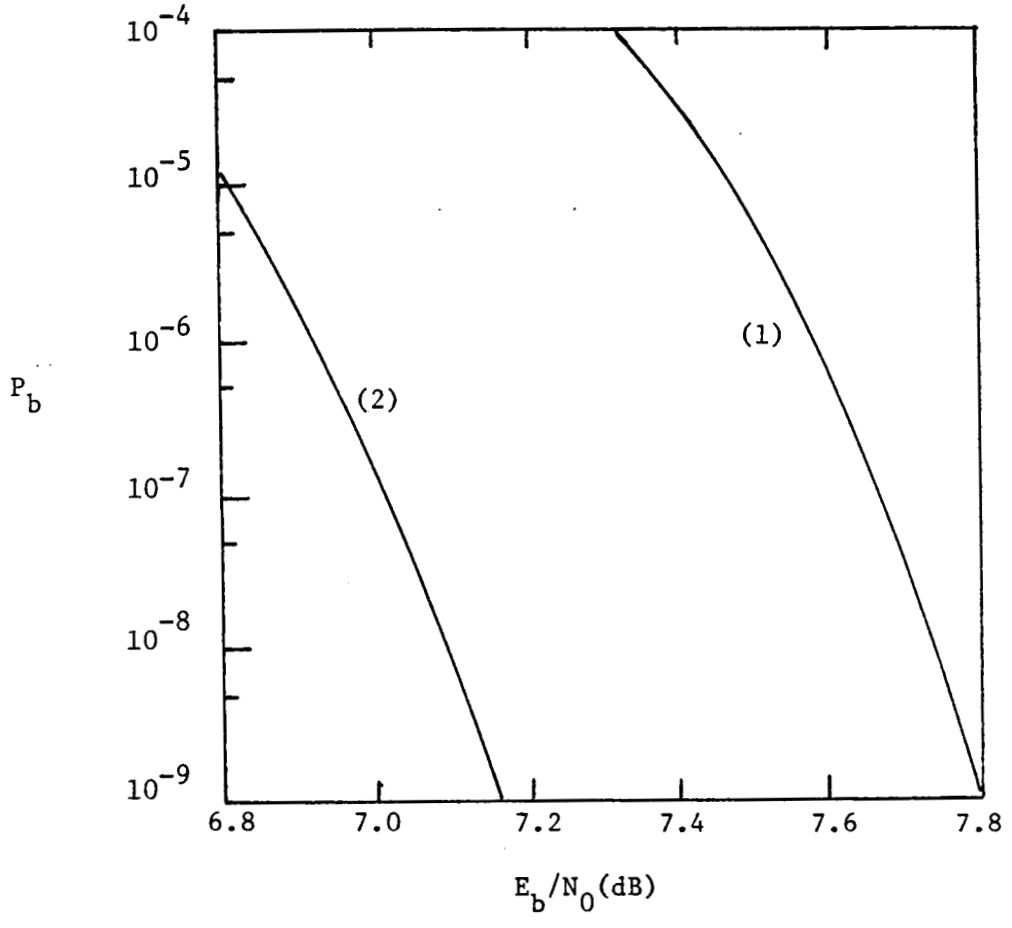


Fig. 9 Performance comparison for concatenated coding systems with

(1) An R_1 = 8/9, 4-state, 2-D PTVTC/8PSK inner code and an (255,201) RS outer code;

(2) An R_1 = 8/9, 4-state, 6-D TC8PSK inner code and the same outer code.

Appendix B

An Expungated Lower Bound on the Free Euclidean Distance of Trellis Codes

AN EXPURGATED LOWER BOUND ON THE FREE EUCLIDEAN DISTANCE OF TRELLIS CODES

MARC ROUANNE
DANIEL J. COSTELLO, Jr.
Dept. of Elec. & Comp. Engr
Univ. of Notre Dame
Notre Dame, IN 46556

ABSTRACT

An expurgated lower bound on the minimum free Euclidean distance d_{free} of Trellis Coded Modulation (TCM), which guarantees the existence of good TCM schemes of any complexity, is presented. Random coding lower bounds (Gilbert bounds) on the free distance of binary codes in particular, and of trellis codes in general, are not very tight for small constraint lengths, whereas they are exponentially tight for large constraint lengths. We tighten the best lower bound known on the free Euclidean distance of TCM schemes by expurgating the set of codes for which the bound holds to obtain an "expurgated bound on d_{free}". The expurgated bound is tighter than any other lower bound for small constraint lengths and is equivalent to the best asymptotic bounds for large constraint lengths.

INTRODUCTION

An efficient binary trellis encoder assigns binary codewords to branches to achieve the largest possible minimum Hamming distance d_h between output sequences generated by different information sequences. This guarantees minimum error probability when using maximum likelihood decoding. Similarly, an efficient non-binary Trellis Coded Modulation (TCM) scheme assigns channel signals to branches to achieve maximum d_{free}, and hence minimum error probability when using maximum likelihood decoding. Many of the best known trellis codes show a rapid increase in the metric (Hamming or Euclidean distance) between diverging

This work was supported by NASA grant NAG5-557-02 and NSF grant ECS84-14608.

and remerging branches in the trellis. In other words, the first and last branches of error events contribute most of the overall distance between a correct path and an error path, at least for short constraint lengths.

A (k,ν) trellis code corresponds to k shift registers of lengths $\nu_i \leq \nu$ $(1 \leq i \leq k)$. k is the input width of the trellis code and ν its memory order. Here, we assume that all the shift registers have the same length ν so that $\nu_0 = k\nu$ is the constraint length of the trellis code. Two branches in a trellis are *adjacent* if they leave or reach the same state (Figure 1). The minimum adjacent distance d_{adj} of a trellis code is the minimum distance between adjacent branches in the trellis. Ungerboeck searched for pseudo-optimum TCM schemes with a large increase in the metric at the ends of error events for 2-dimensional signal constellations [1]. Most of his codes still stand as the best known codes. Wei generalized Ungerboeck's approach, and constructed TCM schemes with large or even maximum d_{adj} for multidimensional signal constellations [2]. In general, most of the TCM schemes presented by Forney achieve a large minimum adjacent distance [3]. All these results indicate that codes with large adjacent distances tend to be good codes on the average. Therefore, we derive an expurgated lower bound for a set of codes with large minimum adjacent distance.

The expurgated lower bound on the free Euclidean distance of trellis codes consists of two parts: a random coding bound on the distance between branches in the middle of error events and a lower bound on the distance between branches at the ends of error events. For small constraint lengths, the major contribution to an error path metric comes from the second part of the bound, whereas for large constraint lengths the first part of the bound is more important.

THE SET OF CODES

The derivation of the expurgated lower bound on the free Euclidean distance of trellis codes is an extension of the derivation of the random coding bound derived for the set of all time-varying Euclidean Trellis Codes [4] which owes much to Chernoff, Gallager, Viterbi and Forney [5] -[8]. The first difference between the two derivations comes from the definition of the set of codes. For any $\delta \geq 0$, C_δ denotes the set of (k,ν) trellis codes c which have minimum adjacent distance $d_{adj} \geq \delta$. For any δ, the expurgated bound lower bounds the average free Euclidean distance of codes in C_δ.

The following Lemmas answer the questions: Is it possible, for the same trellis code, to maximize the distance between diverging branches and the distance between remerging branches? How do these maximizations affect the distances between branches that are not adjacent? The Lemmas are based on the notion of 2-state *subtrellis*. For example, Figure 2b shows the two 2-state subtrellises of the 4-state trellis in Figure 2a.

Lemma 1: Given a signal constellation S which contains M equiprobable signals, Y_t and Y_t' represent two non-adjacent branches which correspond to the same time interval t, and y_t is the signal labeling Y_t, then the probability $p(y_t'|y_t)$ over all codes in C_δ that y_t' is the label for Y_t' is

(i) $p(y_t'|y_t) \leq M\, p(y_t')$ if Y_t and Y_t' belong to the same 2-state subtrellis,

(ii) $p(y_t'|y_t) = p(y_t')$ if Y_t and Y_t' do not belong to the same 2-state subtrellis.

Lemma 2 upper bounds the probability of a particular labeling of an error path given the labeling of the correct path over all codes in C_δ.

Lemma 2: Given a labeled path $\underline{y}$ in a (k, ν) trellis code in C_δ, the probability of a labeling $\underline{y}'$ for an error path of length τ is bounded by

$$p(\underline{y}'|\underline{y}) = \prod_{t=1}^{\tau} p(y_t'|y_t) \leq p(y_1'|y_1)\, p(y_\tau'|y_\tau)\, M^{2B(\tau,\nu)} \prod_{t=2}^{\tau-1} p(y_t'|y_t),$$

where $B(\tau,\nu) \triangleq \lceil \tau/\nu - 2 \rceil_+$ and $\lceil x \rceil_+$ is the smallest non-negative integer larger than or equal to x.

Proving Lemma 2 consists in counting the number of branches in an error event which satisfy (i) or (ii) and are not at the ends of the error event.

DERIVATION OF THE BOUND

The first step consists in stating what must be proved

Assertion: For any $\delta \geq 0$ and any expurgated set C_δ of codes c and for any $\alpha \geq 0$

$$e^{\alpha d^2} \sum_{c \in C_\delta} p(c) \sum_{\underline{y} \in c} \sum_{\underline{y}' \in e_{\underline{y}}(c)} e^{-\alpha d_e^2(\underline{y},\underline{y}')} < 1 \Rightarrow \exists\, c_0 \in C_\delta \text{ s.t. } d_{free}(c_0) > d,$$

where $e_{\underline{y}}(c)$ is the set of error paths in c diverging from $\underline{y}$ at time 1.

The next step, what Forney calls configuration counting, consists of regrouping the labeled paths $\underline{y}$ (or $\underline{y}'$) that correspond to the same unlabeled paths $\underline{Y}$ (or $\underline{Y}'$). This is very similar to configuration counting over the set C of all time varying trellis codes, since it essentially depends on the topology of the trellis. This step depends heavily on Lemma 2. The last step puts the bound in its final form:

Theorem: Given a signal constellation S and a distance δ, there exists a (k, ν) trellis code with maximum minimum adjacent distance $d_{adj} \geq \delta$ and minimum free Euclidean distance d_{free} such that:

$$d_{free}^2 > 2\delta^2 + \max_{\substack{E(\alpha) > k\ln 2 + 1/\nu \ln M \\ 0 \leq \alpha}} \left[(\nu - 2)\frac{E(\alpha)}{\alpha} - \frac{\theta'[E(\alpha)]}{\alpha} \right],$$

where

$$E(\alpha) \triangleq -\ln \sum_{y \in S} p(y) \sum_{y' \in S} p(y')\, e^{-\alpha d_e^2(y,y')},$$

$d_e(y,y')$ is the Euclidean distance between signals y and y',

$$\theta'[E(\alpha)] \triangleq \ln \frac{1}{e^{\eta} - 1} \cdot \frac{1 - e^{-\nu\eta}}{1 - e^{-\nu\eta + \ln M}},$$

$\eta \triangleq [E(\alpha) - k\ln 2]$, and

α is a parameter which optimizes the bound.

APPLICATION OF THE BOUND

The expurgated bound is mostly useful for small constraint lengths, since it is equivalent to random coding bounds for large constraint lengths. For example, if the Euclidean distance can be made proportional to the Hamming distance as in BPSK modulation, then the bound applies to binary trellis codes like convolutional codes. Table 1 shows a comparison of the expurgated bound with the best random coding bound [9], the best known codes [10], and Heller's very tight upper bound [11], for rate 1/4 convolutional codes.

Since upper bounds are not as tight as the Heller bound for TCM schemes used with M-ary modulation, the bound is most useful for trellis codes used with M-ary modulation like M-PAM or M-PSK signal constellations. Figure 3 shows the expurgated bound for 8-PSK modulation compared with the best known rate 2/3 codes [1], upper bounds [12]-[13], and the random coding bound [4].

The expurgated bound, like any other random coding bound on convolutional or trellis codes, applies only to rate k/n codes generated by k shift registers of equal length. However, trellises that represent shift registers of various lengths can be decomposed into *nested trellises* [14], to which the bound applies. Figure 4 shows a general representation of nested trellises. Figure 5 shows a 2 level nested trellis (Figure 5c) where the branches of the trellis in Figure 5b have been replaced with the trellis in Figure 5b (the input width k determines the number of levels of a nested trellis). For example, Ungerboeck's trellises are nested trellises where parallel branches represent 1-state trellises (Figure 5d). When the distance between parallel branches limits the achievable free distance of a code, they can be replaced with trellises to introduce redundancy between sequences of signals from the same subsets (Figure 4) [15]. The free distance of a nested trellis code satisfies

$$d_{free} = \min_{1 \le i \le k} (d_{free}(\nu_i)),$$

where $d_{free}(\nu_i)$ is the free distance of any subtrellis of level i, where a subtrellis of level i represents the last $k - i + 1$ shift registers [14]. The expurgated bound on d_{free} can be used to lower bound each $d_{free}(\nu_i)$ by replacing the distances between signals with distances between subsets of signals that correspond to level i (Figure 4). This gives a lower bound on the minimum free Euclidean distance which applies to any trellis code.

A more general and tighter version of the bound can be derived, but involves a more detailed proof [15]. In general, this bound shows that random coding bounds can be improved significantly by restricting the set of codes. It also shows that codes with a large increase in the metric at the ends of error events perform well on the average. Finally, it can be used to compare TCM schemes with small constraint lengths without having to search for the corresponding codes.

REFERENCES

[1] G. Ungerboeck, "Channel Coding with Multilevel Phase Signals", IEEE Trans. Inform. Theory, vol. IT-28, pp. 55-67, Jan. 1982.

[2] L.F. Wei, "Trellis-Coded Modulation with Multi-dimensional Constellations", IEEE Trans. Inform. Theory, Vol. IT-33, pp. 483-501, July 1987.

[3] G.D. Forney, Jr., "Coset Codes I: Geometry and Classification", IEEE Trans. Inform. Theory, to appear.

[4] M.F. Rouanne and D.J. Costello, Jr., "A Lower Bound on the Minimum Euclidean Distance of Trellis Coded Modulation Schemes", IEEE Trans. Inform. Theory, to appear.

[5] H. Chernoff, "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on a Sum of Observations," Ann. Math. Stat. , vol. 23, pp. 493-507, 1952.

[6] R. G. Gallager, "A simple Derivation of the Coding Theorem and Some Applications," IEEE Trans. Inform. Theory , vol. IT-11, pp. 3-18, Jan. 1965.

[7] A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," IEEE Trans. Inform. Theory , vol. IT-13, pp. 260-269, Apr. 1967.

[8] G. D. Forney, Jr. "Convolutional Codes II: Maximum Likelihood Decoding", Infor- mation and Control , vol. 25, pp. 222-266, July 1974.

[9] D. J. Costello, Jr. "Free Distance Bounds for Convolutional Codes", IEEE Trans. Inform. Theory, vol. IT-20, pp. 356-365, May 1974.

[10] K. J. Larsen, "Short convolutional codes with maximal free distance for rates 1/2, 1/3, and 1/4," IEEE Trans. Inform. Theory, vol. IT-19, pp. 371-372, May 1973.

[11] J.A. Heller, "Short constraint length convolutional codes," Jet Propul. Lab., Calif. Inst. Tech., Pasadena, CA. Space Programs Summary 37-54, vol. 3, pp. 171-174, Dec 1968.

[12] G.J. Pottie and D.P. Taylor, "Sphere-Packing Upper Bounds on the Free Distance of Trellis Codes", IEEE Trans. Inform. Theory, to appear.

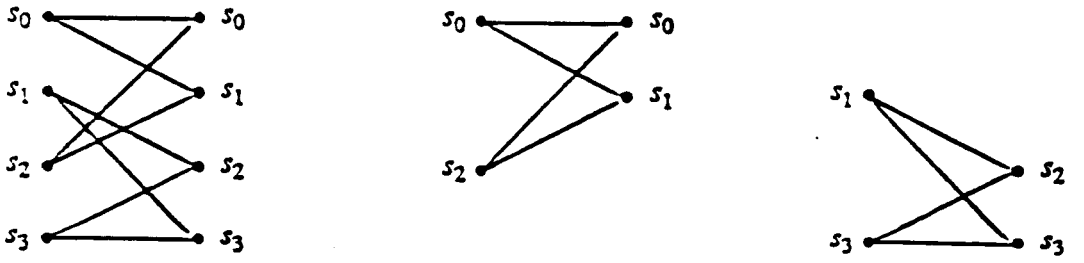
[13] M. Burnashev and E. Biglieri, "Bounds on the Minimum Distance of Trellis Codes", submitted to IEEE Trans. Inform. Theory, 1986.

[14] M.F. Rouanne and D.J. Costello, Jr., "Nested Trellis Codes", in preparation.

[15] M.F. Rouanne and D.J. Costello, Jr., "An Expurgated Lower Bound on the Free Distance of Trellis Coded Modulation Schemes", in preparation.

(a) 1 input bit (b) 2 input bits

Figure 1: Adjacent distances for 1 or 2 input bits.



4-STATE TRELLIS
(a)

2 2-STATE SUBTRELLISES
(b)

Figure 2: 2-state subtrellises in a 4-state trellis.

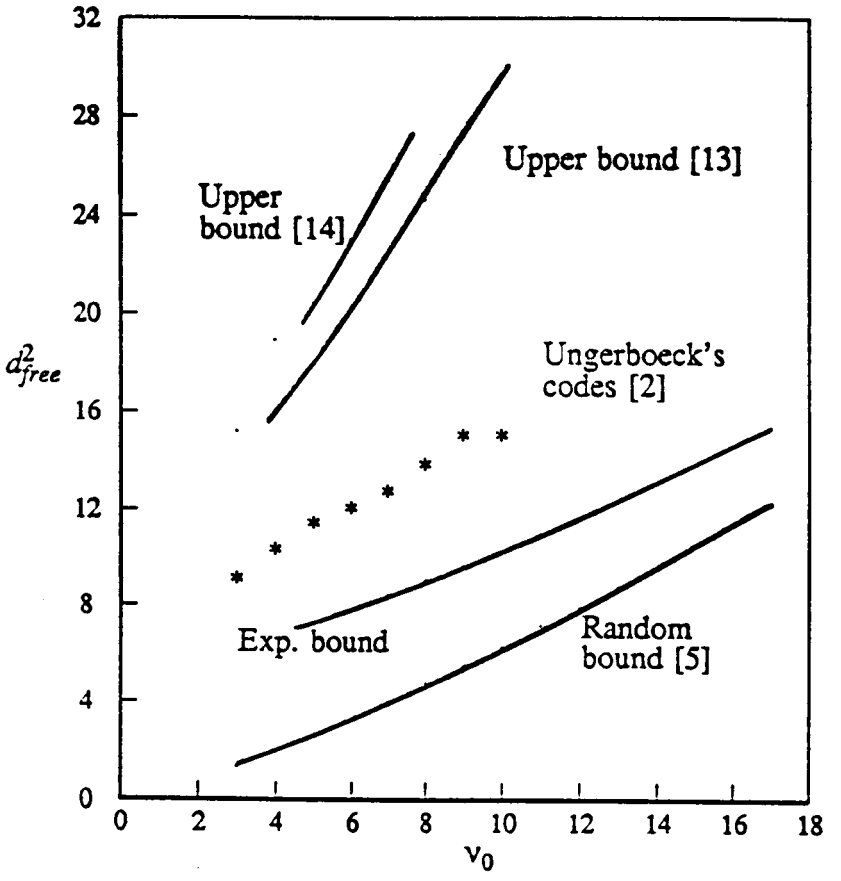


Figure 3: Bounds and on the free Euclidean distance of trellis coded 8-PSK.

Table 1: Bounds on the minimum free Hamming distance of trellis coded BPSK

d_k for rate 1/4				
ν_0	rand. bound	exp. bound	codes	upper bound
2	4	9	10	10
3	5	10	13	13
4	7	12	16	16
5	8	13	18	18
6	10	15	20	20
7	12	16	22	22
8	13	18	24	24
9	15	19	27	27
10	16	21	29	29
11	18	23	32	32
12	20	24	33	33
13	21	26	36	36

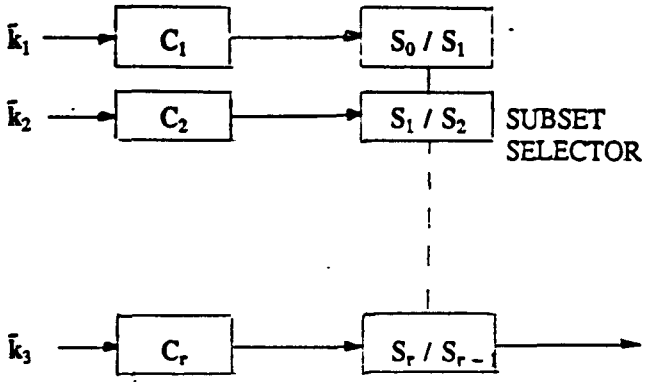
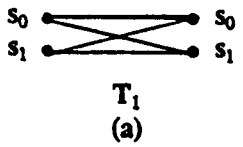
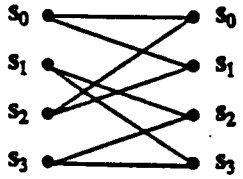


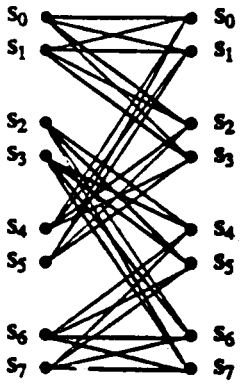
Figure 4: A generalized Trellis Coded Modulation scheme (nested trellis).



T_1
(a)

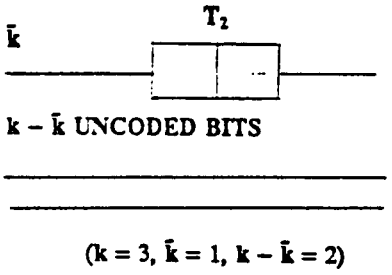


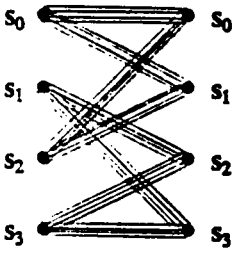
T_2
(b)



NESTED TRELLIS T_2T_1.
(c)

BRANCHES OF T_2 ARE REPLACED WITH T_1 TO OBTAIN T_2T_1.





$T_2T_0^3$

EXAMPLE OF UNGERBOECK'S TRELLIS CODE.
(d)

Figure 5: Examples of nested trellises.

Appendix C

Bandwidth Efficient Coding on Channels With Phase Jitter

BANDWIDTH EFFICIENT CODING ON CHANNELS WITH PHASE JITTER [1]

Christian Schlegel and Daniel J. Costello, Jr.
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

Abstract

Bandwidth efficient Trellis Code Modulation (TCM) is usually designed to maximize the free Euclidean distance d_{free}. While this assures good performance on AWGN channels it is shown that such codes suffer severe degradation on channels with phase errors. The problem is compounded due to the fact that the performance gain of TCM coding schemes is achieved at the expense of expanding the signal sets used, which results in poorer phase synchronization.

A generalized distance measure is introduced which becomes the new design criterion for coding on channels with phase jitter. This generalized distance is found by applying the Chernoff bounding technique to these channels. The performance of existing codes is compared to some new codes, designed with the new criterion in mind. Hard quantizing the output of the jitter channel in conjunction with RS-codes in order to avoid dealing with unreliable distance metrics is also examined.

[1]This work was supported by NASA Grant NAG5-557 and NSF Grant ECS 84-14608

BANDWIDTH EFFICIENT CODING ON CHANNELS WITH PHASE JITTER [1]

Christian Schlegel and Daniel J. Costello, Jr.
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

Summary

In recent years bandwidth efficient Trellis Coded Modulation (TCM) with multilevel/phase signaling introduced by Ungerboeck [1] has become increasingly popular and much analysis has been devoted to the performance of these coding schemes on AWGN channels. In this paper we examine AWGN channels that further suffer from phase jitter due to synchronization inaccuracies of the phase recovery loop. This phase jitter materializes as a time varying phase error $\theta(t)$ by which the signal constellation is rotated. We assume throughout the paper that DSB AM-modulation is used with coherent reception and that the phase error therefore affects two dimensions at a time (i.e., the quadrature and in-phase components). We further assume at each signaling interval that the receiver has no information about the phase error, because if it had such information, the received 2-dimensional signals could simply be rotated back by the erroneous phase angle into the correct position.

An upper bound is presented on the two code word error probability for coded communication systems on an arbitrary memoryless channel. This bound, which is based on the Chernoff bounding technique, has the form

$$P_2[\mathbf{x} \rightarrow \hat{\mathbf{x}}] \leq \prod_{j=1}^{N} C(\hat{\underline{x}}_j, \underline{x}_j), \tag{1}$$

where $\mathbf{x} = [\underline{x}_1, \cdots \underline{x}_N]$ and $\hat{\mathbf{x}} = [\hat{\underline{x}}_1, \cdots \hat{\underline{x}}_N]$ are the two code words of length N and $C(\hat{\underline{x}}_j, \underline{x}_j)$ is the Chernoff factor of the two 2-dimensional signals $\underline{x}_j$ and $\hat{\underline{x}}_j$. Using the bound (1) on the two code word error probability, the cutoff rate R_0 for a random ensemble of codes and the transfer function bound for specific codes can be evaluated immediately.

We then apply (1) to the "phase jitter" channel. The carrier phase is usually tracked with a first-order phase locked loop as discussed in [2] and [3]. The phase error $\theta(t)$ in the steady state is then known to obey a Tikhonov density function which can be closely approximated by the Gaussian density function for phase deviations σ_θ as large as 10 degrees. Other contributions to the phase error can also be modeled as Gaussian processes, as discussed in [4]. We therefore use the Gaussian denisty function for θ in the analysis, i.e.,

[1]This work was supported by NASA Grant NAG5-557 and NSF Grant ECS 84-14608

$$p(\theta) \approx \frac{1}{\sqrt{2\pi\sigma_\theta^2}} e^{-\theta^2/2\sigma_\theta^2}. \tag{2}$$

We then show that for the phase jitter channel the Chernoff factor becomes

$$C(\underline{x}_j, \hat{\underline{x}}_j) \approx e^{-\frac{1}{4N_0}\left\{|\underline{x}_j - \hat{\underline{x}}_j|^2 - \frac{2\rho|\underline{x}_j|^2|\hat{\underline{x}}_j|^2 \sin^2(\theta_j)}{1-\rho(|\underline{x}_j|^2 - |\underline{x}_j||\hat{\underline{x}}_j|\cos(\theta_j))}\right\}}, \tag{3}$$

where $\rho = \sigma_\theta^2/N_0$ is the ratio of phase noise power to channel noise power. The above equation shows that in the case of phase jitter a new distance measure dominates the performance of a code. This distance measure

$$d(\mathbf{x}, \hat{\mathbf{x}}) = \sum_{j=1}^{N} \left(|\underline{x}_j - \hat{\underline{x}}_j|^2 - \frac{2\rho|\underline{x}_j|^2|\hat{\underline{x}}_j|^2 \sin^2(\theta_j)}{1 - \rho(|\underline{x}_j|^2 - |\underline{x}_j||\hat{\underline{x}}_j|\cos(\theta_j))} \right) \tag{4}$$

is still additive and reduces to the Euclidean distance for negligible phase noise, i.e., for $\rho \to 0$.

We contrast the performance of trellis codes with maximum Euclidean distance to TCM codes specifically designed for phase jitter channels. For example, a 64-state Trellis Code using 8-PSK signaling with a gain of 4.77 dB on an AWGN channel suffers a degradation of 0.38 dB for $\rho = 0.1$, 1.25dB for $\rho = 0.2$, and 5.07dB on a severly impaired channel with $\rho = 0.4$. Another bandwidth efficient coding scheme employs Reed Solomon block codes with multilevel/phase signaling in conjunction with hard decision decoding. This method is inferior to TCM coding on AWGN channels and for small SNR values, but becomes more attractive on channels with phase jitter because the hard decision decoding is less sensitive to phase errors.

References

[1] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," *IEEE Trans. Inform. Theory*, Vol. IT-28, No.1, pp. 55–67, January, 1982.

[2] T. L. Lim, "Error Probability Bounds for Unbalanced QAM Over the Gaussian Phase Jittered, Dispersive Channel", *IEEE Trans. Comm.*, Vol. COM-30, No.7, July 1982, pp. 1729–1735.

[3] A. J. Viterbi, *Principles of Coherent Communications*, New York, McGraw-Hill, 1966.

[4] H. Leib and S. Pasupathy, "Trellis-Coded MPSK with Reference Phase Errors", *IEEE Trans. Comm.*, Vol. COM-35, No.9, September 1987, pp. 888–900.

Appendix D

Construction of Good Trellis Codes

CONSTRUCTION OF GOOD TRELLIS CODES

MARC ROUANNE
DANIEL J. COSTELLO, Jr.
Dept. of Elec. & Comp. Engr
Univ. of Notre Dame
Notre Dame, IN 46556

ABSTRACT: We present a construction technique for good trellis codes with large constraint length. The technique can be used for any practical constraint length, any signal constellation, and any code rate. The construction does not require an exhaustive search, so that the complexity of the construction does not increase significantly with the constraint length. However, it requires tables of optimum distance profile rate 1/m trellis codes for small constraint lengths. Rate k/n trellis codes are constructed from a nesting of good rate 1/m trellis codes ($m \leq n$). The trellis codes constructed are regular and the construction technique also applies to convolutional codes. The codes obtained are generally slightly sub-optimum, but in certain instances, the codes achieve an optimum distance profile and a maximum free distance. The construction technique is interesting both from a theoretical and a practical point of view. To the author's knowledge, it is the first systematic construction technique for trellis codes which can be used with any rate, constraint length, and modulation scheme, and it gives some insight into the structure of trellis and convolutional codes. From a practical point of view, good distance profile codes with large constraint lengths will perform well with sequential decoding, where it is not so crucial to have optimum free distance codes.

This work was supported by NASA grant NAG5-557 and NSF grant ECS84-14608.

CONSTRUCTION OF GOOD TRELLIS CODES

MARC ROUANNE
DANIEL J. COSTELLO, Jr.
Dept. of Elec. & Comp. Engr
Univ. of Notre Dame
Notre Dame, IN 46556

SUMMARY

We present a construction technique for good trellis codes with large constraint length. The technique can be used with any practical constraint length, any signal constellation, and any code rate. The construction does not require an exhaustive search, so that the complexity of the construction does not increase significantly with the constraint length. However, it requires tables of optimum distance profile rate 1/m trellis codes for small constraint lengths. The trellis codes constructed are regular and the construction technique also applies to convolutional codes. Once the code is constructed we check its distance profile. We generally obtain slightly sub-optimum codes, but in certain instances the codes achieve an optimum distance profile and a maximum free distance.

The construction technique is interesting both from a theoretical and a practical point of view. To the author's knowledge, it is the first systematic construction technique for trellis codes which can be used with any rate, constraint length, and modulation scheme, and it gives some insight into the structure of trellis and convolutional codes. From a practical point of view, good distance profile codes with large constraint lengths will perform well with sequential decoding, where it is not so crucial to have optimum free distance codes.

The construction technique nests rate 1/m trellis codes to obtain rate k/n trellis codes ($m \leq n$). Note that some rate 1/m trellis codes can be replaced with block codes to obtain most of the previously known trellis codes [1]-[3]. First, the signal constellation is partitioned into nested subsets. These nested subsets are then used as the signal constellation for the nested trellis codes, like subsets are used for uncoded bits in some

This work was supported by NASA grant NAG5-557 and NSF grant ECS84-14608.

known trellis codes [1]. From tables of optimum rate 1/m trellis codes, we choose an optimum distance profile code for each level of the signal constellation partitioning. These rate 1/m codes must be good so that the upper bound on the free distance of the nested trellis codes achieves a large value [4]. Then, the rate 1/m subtrellises can be nested to form a rate k/n trellis code. Finally, the distance profile of the rate k/n trellis code can be computed. So far, there is no theoretical lower bound on the free distance of nested trellis codes, so that it cannot be guaranteed that the resulting code is good. However, we always obtained good codes by applying this technique or slightly modified versions of it [5].

The following table shows some rate 2/3 binary codes obtained through the above construction. We do not include codes of large constraint length here, although they can be easily constructed. Each code consists of two rate 1/2 nested trellis codes. The rate 1/2 codes are optimum binary convolutional codes. The table shows the achieved Hamming free distance, the free distance of optimum codes and of Optimum Distance Profile (ODP) codes, and the upper bound on the free distance of the nested trellis codes. For simplicity, we chose rate 1/2 sub-trellises, although rate 1/3 sub-trellises give a larger value of the upper bound, and generate some better nested codes.

References:

[1] G. Ungerboeck, "Channel Coding with Multilevel Phase Signals", IEEE Trans. Inform. Theory, vol. IT-28, pp. 55-67, Jan. 1982.

[2] L.F. Wei, "Trellis-Coded Modulation with Multi-dimensional Constellations", IEEE Trans. Inform. Theory, Vol. IT-33, pp. 483-501, July 1987.

[3] G.D. Forney, Jr., "Coset Codes I: Geometry and Classification", IEEE Trans. Inform. Theory, to appear.

[4] M. Rouanne and D.J. Costello, Jr., "An Expurgated Lower Bound on the Minimum Euclidean Distance of Trellis Codes", Presented at Allerton Conf. on Commun., Cont., and Comp., Monticello, IL, October 1987.

[5] M. Rouanne and D.J. Costello, Jr., "Nested Trellis Codes", in preparation.

memory order	constraint length	generator sequences			d_{free}	max[1] d_{free}	ODP[2] d_{free}	bound[3] d_{free}
1	2	3 0	1 3	0 1	3	3	3	3
2	3	1 0	3 5	0 7	3	4	4	3
2	4	5 0	7 5	0 7	5	5	5	5
3	5	15 0	17 17	0 15	5	6	6	5
3	6	13 0	17 13	0 17	6	7	6	6
4	7	17 0	13 31	0 27	6	8	7	6
4	8	23 0	35 23	0 35	7	8	8	7
5	9	31 0	27 53	0 57	7	9	8	7
5	10	75 0	53 75	0 53	8	10	9	8
6	11	75 0	53 163	0 135	8	-	10	8
6	12	163 0	135 163	0 135	10	-	10	10

1 Optimum rate 2/3 binary codes.

2 Optimum distance profile codes.

3 Upper bound on the free distance of the nested trellises constructed from rate 1/2 trellis codes.

Appendix E

Parity Retransmission Hybrid ARQ Using Rate 1/2 Convolutional Codes on a Non-Stationary Channel

Parity Retransmission Hybrid ARQ Using Rate 1/2 Convolutional Codes on a Non-Stationary Channel

Laurent R. Lugand

French Institute for Iron and Steel
Computer Department
78105 Saint Germain en Laye, FRANCE

Daniel J. Costello, Jr.
Robert H. Deng

Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, IN 46556

ABSTRACT

Hybrid Automatic-Repeat-Request (ARQ) error control coding makes use of both error detection and error correction in order to achieve high throughputs and low undetected error probabilities on two-way channels. For non-stationary channels, where the channel bit error rate (BER) varies over time, the technique of parity retransmission allows the error control strategy to adapt to the state of the channel.

In this paper we propose a parity retransmission hybrid ARQ scheme which uses rate 1/2 convolutional codes and Viterbi decoding. The performance analysis is based on a two-state Markov model of a non-stationary channel. Throughput efficiency is shown to improve as the channel becomes more bursty in nature.

July 30, 1987

This work was supported by the National Aeronautics and Space Administration under grant NAG 5-557 and by the National Science Foundation under grant ECS 84-14608.

Parity Retransmission Hybrid ARQ Using Rate 1/2 Convolutional Codes on a Non-Stationary Channel

Laurent R. Lugand

French Institute for Iron and Steel
Computer Department
78105 Saint Germain en Laye, FRANCE

Daniel J. Costello, Jr.
Robert H. Deng

Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, IN 46556

1. Introduction

Automatic–repeat–request (ARQ) strategies have long been utilized to control errors on two-way digital transmission links. Most of the work in this area has been done using block codes with error detection only, due to the packetized nature of the messages and the relatively low coding overhead allowed in many systems. However, in systems where the packet lengths are relatively large, say on the order of 1000 bits or more, and where the noise and/or interference levels are high, error detection only results in a low throughput due to the large number of retransmissions required. Satellite networks [1] and packet radio [2] are examples of such systems. In these instances a combination of error correction and error detection can offer significant advantages over an error detection only system. This is called *hybrid ARQ error control.*

Two basic types of hybrid ARQ error control strategies have been considered. The first type includes parity bits for both error detection and error correction in each transmitted packet. The decoder will correct those received packets within the error-correcting capability of the code, while requesting a retransmission of those packets with detectable but uncorrectable errors. Since bits for error correction are sent with every packet, the code rate places an upper limit on the throughput efficiency of the system. For this reason, this strategy is best suited for systems in which a fairly

This work was supported by the National Aeronautics and Space Administration under grant NAG 5-557 and by the National Science Foundation under grant ECS 84-14608.

constant level of noise and interference is anticipated on the channel. In this case, enough error correction overhead can be designed into the system to correct the vast majority of packets, thereby greatly reducing the number of retransmissions compared to an error detection only scheme and enhancing the system throughput. On the other hand, if the channel is quiet most of the time and noisy only on occasion (a *non-stationary channel*), designing a code to correct the occasional noisy bursts will reduce throughput compared to an error detection only scheme because the error correction overhead is wasted during the quiet periods.

In the second type of strategy, bits for error detection only are sent on the first transmission. If errors are detected and a retransmission requested, parity bits on the original information packet are sent along with some bits for error detection. If no errors are detected on the second transmission, the parity bits are inverted to recover the original information. If errors are detected, the two received packets are treated together as a code word in a rate 1/2 code. If the error-correcting-capability of the code is exceeded, and decoding is unreliable, the original transmission is repeated. This process continues, alternating transmissions between the original data packet and the parity packet, until either an error-free packet is received or error correction is possible. This strategy is referred to as *parity retransmission* [3].

Since parity bits for error detection only are sent on the first transmission, the upper limit on throughput efficiency is near 1. Throughput suffers only when retransmissions are required, since it is only then that parity bits for error correction are sent. In other words, parity bits for error correction are transmitted only when they are needed. It is this feature which gives the parity retransmission strategy the ability to adapt to changing channel conditions. When the channel is quiet, parity bits for error detection only are transmitted, and a high throughput is maintained. Only when noise or interference causes packets to be received incorrectly are parity bits for error correction transmitted, resulting in a reduced throughput. This adaptive capability of the parity retransmission strategy is particularly useful in applications such as satellite communication and packet radio, where fluctuating channel conditions due to fading and interference are commonly encountered.

Either block or convolutional codes can be used with both types of strategy. Schemes using the first strategy and block codes have been in existence for quite some

time [4-9]. Several of these schemes using convolutional codes have also appeared in the literature [2,10-12]. Parity retransmission using block codes was introduced more recently [3,13,14]. Several similar schemes using convolutional codes have also been presented [15-18]. Although all the parity retransmission schemes that have appeared in the litreature have been proposed for use on non-stationary channels, the analysis in each case used a stationary channel model.

In this paper, the performance of a parity retransmission hybrid ARQ scheme using rate 1/2 convolutional codes on a non-stationary channel is analyzed. In section 2, a protocol is described which is capable of achieving higher throughputs than previously proposed parity retransmission schemes. In section 3, a two-state Markov chain channel model is defined. This model constitutes a first approximation to a non-stationary channel. In sections 4 and 5, the two-state channel model is used to analyze the throughput and undetected error probability of the protocol presented in section 2, when the receiver has both an infinite and a finite buffer size. It is shown that the throughput improves as the channel becomes more bursty. This corresponds with our intuitive notion that parity retransmission schemes are best suited for non-stationary or bursty channels. In section 6, performance curves are calculated for a particular example.

2. Description of the Protocol

The parity retransmission hybrid ARQ scheme with a rate 1/2 convolutional code employs two codes, C_0 and C_1. C_1 is a (2, 1, m) convolutional code with two generator polynomials, $G_1(x)$ and $G_2(x)$, and it is used for error detection and error correction. C_0 is a high rate ($n-m$, $n-m-r$) binary block code used for error detection only, where r is the number of parity bits in C_0.

When an ($n-m-r$)-bit information sequence $I(x)$ is generated, it is first encoded into a code vector, denoted by $J(x)$, in the ($n-m$, $n-m-r$) block code C_0. Then the ($n-m$)-bit vector $J(x)$ is encoded into an n-bit code vector $V_1(x) = J(x) \cdot G_1(x)$, and $V_1(x)$ is transmitted over the channel.

Let $\hat{V}_1(x)$ be the noisy version of $V_1(x)$ arriving at the receiver. The syndrome of $\hat{V}_1(x)$ is checked in two steps. First $\hat{V}_1(x)$ is considered as a noisy version of a code word in the (n, $n-m$) shortened cyclic code generated by $G_1(x)$. The syndrome of the

shortened cyclic code is checked; if it is zero, we have an estimate $\hat{J}(x)$ of $J(x)$. Next the syndrome of $\hat{J}(x)$ in the high rate $(n-m, n-m-r)$ block code is checked; if it is zero, we have an estimate $\hat{I}(x)$ of $I(x)$. The estimate $\hat{I}(x)$ is assumed to be error-free and is delivered to the data sink. In this case, we call $\hat{V}_1(x)$ a *zero syndrome vector* (ZSV). If, however, the first or the second syndrome check is negative, i.e., $\hat{V}_1(x)$ is a *nonzero syndrome vector* (NSV), then a NACK signal is sent to the transmitter and $\hat{V}_1(x)$ is stored in a receiver buffer. The transmitter then sends a second vector $V_2(x) = J(x) \cdot G_2(x)$, and $\hat{V}_2(x)$ is received after a round trip delay. Its syndrome is checked in two steps in the same way as $\hat{V}_1(x)$. If both syndromes are zero, $\hat{V}_2(x)$ is assumed to be error-free and an estimate $\hat{I}(x)$ of $I(x)$ is recovered directly; otherwise, $\hat{V}_1(x)$ and $\hat{V}_2(x)$ are decoded using the Viterbi algorithm, producing an estimate $\hat{J}(x)$. The syndrome of $\hat{J}(x)$ is then checked using the $(n-m, n-m-r)$ block code. If it is zero, $\hat{I}(x)$ is recovered and delivered to the data sink; if it is nonzero, a second NACK is sent requesting the retransmission of $V_1(x)$. The previously received version of $V_1(x)$, $\hat{V}_1(x)$, is discarded at the receiver and replaced by the new one. The receiver continues checking the syndrome of each received vector, trying to decode using the Viterbi algorithm if the syndrome check is negative, and requesting a retransmission of $V_1(x)$ or $V_2(x)$ in alternating order if the decoding is unsuccessful, until the information vector is delivered to the data sink.

3. Description of the Channel Model

Two-state model

Let us model the channel as a Markov chain (see Fig.1). State 0 is the *quiet* state, where the bit error rate (BER) is ε_0. State 1 is the *noisy* state, where the BER is $\varepsilon_1 \gg \varepsilon_0$. p is the transition probability from state 0 to state 1 and p' is the transition probability from state 1 to itself. To simplify the model's treatment, we assume that one time frame in the model corresponds to the transmission of one data vector, i.e., the noisy bursts last for a multiple of the transmission time of a data vector. This type of model was first introduced by Gilbert [19].

It can be shown that *the average burst length* is

$$\bar{b} = \frac{1}{1 - p'} , \tag{1}$$

the average BER is

$$\overline{\varepsilon} = \frac{(1 - p')\,\varepsilon_0 + p\,\varepsilon_1}{1 - p' + p}\,, \tag{2}$$

and the *duty cycle* of the noisy bursts, or the probability of being in the noisy state , is

$$p_1 = \frac{p}{1 - p' + p}\,, \tag{3}$$

or

$$p_1 = \frac{\overline{\varepsilon} - \varepsilon_0}{\varepsilon_1 - \varepsilon_0}\,. \tag{4}$$

Four parameters govern the two-state channel model. They can be chosen to be $\overline{b}$, $\overline{\varepsilon}$, p_1, and the high-to-low BER ratio $\rho \triangleq \frac{\varepsilon_1}{\varepsilon_0}$ [20].

Burst noise model

Of the four parameters selected, one is the average channel BER and the other three characterize the burstiness of the channel. We shall reduce the number of degrees of freedom by proposing a model for the noise bursts which can be *dense* (low duty cycle p_1 and high intensity, i.e., large high-to-low BER ratio ρ) or *diffuse* (large duty cycle and low intensity), as shown in Fig. 2. These terms were first introduced by Massey [21]. The conditions that we impose on a burst channel model are:

$$\lim_{p_1 \to 0} \varepsilon_0 = 0, \tag{5.1}$$

$$\lim_{p_1 \to 0} \varepsilon_1 = 1/2, \tag{5.2}$$

and

$$\lim_{p_1 \to 1} \varepsilon_0 = \overline{\varepsilon}, \tag{6.1}$$

$$\lim_{p_1 \to 1} \varepsilon_1 = \overline{\varepsilon}. \tag{6.2}$$

Conditions (5) represent the limiting case of a dense burst channel, i.e., $p_1 \to 0$ and $\rho \to \infty$, while conditions (6) represent the limiting case of a diffuse burst channel, i.e., $p_1 \to 1$ and $\rho \to 1$, which is equivalent to a *binary symmetric channel* (BSC).

The two-state channel model described by (1) - (4) does not meet these conditions. In fact, from (4) we see that only condition (6.2) is satisfied. We now modify the two-state channel model such that conditions (5) and (6) are met.

Let

$$\varepsilon_0 = \bar{\varepsilon}\, p_1. \tag{7}$$

From (4) and (7), we have

$$\varepsilon_1 = \frac{\bar{\varepsilon}}{p_1} - (1-p_1)\,\bar{\varepsilon}, \qquad \text{for } p_1 \geq \frac{\bar{\varepsilon} + 1/2 - \sqrt{(1/2-\bar{\varepsilon})(1/2+3\bar{\varepsilon})}}{2\,\bar{\varepsilon}}, \tag{8.1}$$

$$\varepsilon_1 = \frac{1}{2}, \qquad \text{otherwise.} \tag{8.2}$$

From (7), we see that (5.1) and (6.1) are satisfied. From (8.1) and (8.2), we see that (5.2) and (6.2) are also satisfied. Now the burst channel model is completely described by $\bar{\varepsilon}$, p_1, and $\bar{b}$, for if these three parameters are known, p', p, ε_0, and ε_1 can be determined from (1), (3), (7), and (8). Before leaving this section, we note that the special case $p_1 = p = p'$ corresponds to the two-state *block interference* (BI) channel model proposed by McEliece and Stark [22]. The BI channel model is completely determined by p_1 and $\bar{\varepsilon}$.

4. System Throughput and Undetected Error Probability Analysis with an Infinite Receiver Buffer

In the next two sections, we analyze the throughput and the undetected error probability of the parity retransmission hybrid ARQ scheme in the *selective-repeat* mode for both an infinite receiver buffer and a finite receiver buffer. Our analysis is based on the assumption that the feedback channel is error-free. In order to carry out the analysis, we first model the receiver's decoding status as a Markov chain.

Receiver's decoding status

Consider a Markov chain with N possible states: $1, 2, \ldots, N$, where changes in states can occur only at discrete times $t_1, t_2, \ldots, t_n, \ldots$. Let us denote the transition probability from state i to state j after k time units by $p_{ij}(k)$. $p_{ij}(k)$ is called the *k-step transition probability*, and it can be determined from the one-step transition

probabilities, namely, $p_{ij}(1)$, or simply p_{ij}, between all pairs of states. These transition probabilities can be summarized by an $N \times N$ matrix, called the one-step *transition probability matrix:*

$$\underline{P} = [p_{ij}] = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,N} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,N} \\ \vdots & & & \\ p_{N,1} & p_{N,2} & \cdots & p_{N,N} \end{bmatrix} . \tag{9}$$

The probabilities in each row of $\underline{P}$ add to 1. The k-step transition probabilities can then be determined from the k-step transition probability matrix, which is given by

$$[\, p_{ij}(k) \,] = \underline{P}^k . \tag{10}$$

The states in the Markov chain can be divided into several categories. The one of most interest to us is the *absorbing state.* A state i in a Markov chain is an absorbing state if the k-step transition probability $p_{ii}(k) = 1, \quad k = 1, 2, \ldots$. Thus, an absorbing state is a state that cannot reach any other state in the chain except itself. A Markov chain may contain more than one absorbing state.

For the parity retransmission hybrid ARQ scheme described in Section 2, decoding is said to be *successful* if the decoded information vector is accepted by the user (it may be decoded correctly or contain undetected errors). Suppose that an initial NSV has been received. Then the receiver's decoding status can be modeled by the Markov chain shown in Fig.3. In Fig.3, states 00, 01, 10, and 11 mean that the decoder has received two NSVs, sent while the channel was in states 0 and 0, 0 and 1, etc., and that decoding has been unsuccessful. State e indicates an undetected decoding error. State c corresponds to error-free decoding. Once the system is in state e or state c, successful decoding results and retransmissions will therefore terminate. Hence, both state e and state c are absorbing states and we assign $p_{e,e} = p_{c,c} = 1$. On the other hand, if , upon the reception of a retransmission , the decoder cannot recover the data vector, retransmissions will continue (the Markov chain will remain in state 00, 01, 10, or 11) until successful decoding occurs (the Markov chain reaches state e or state c).

To derive the transition probabilities of the Markov chain, we first consider the undetected error probability. Undetected errors occur at two levels: the syndrome check before convolutional decoding and the syndrome check after convolutional

decoding. First we consider the former case. Let Q_i, $i = 0,1$, be the probability of receiving a ZSV with an undetected error pattern while the channel is in state i. Q_i can be upper bounded by [23]

$$Q_i \le 2^{-(m+r)}[1 - 2(1 - \varepsilon_i)^n + (1 - 2\varepsilon_i)^n] , \tag{11}$$

where $m + r$ is the sum of the memory order of the convolutional code and of the number of parity checks in the block code (the two codes are treated at this stage as a single block code with $m + r$ parity checks).

To find the probability of undetected error after convolutional decoding, we need to evaluate the BER at the output of the convolutional decoder. The state ij in Fig. 3 means that a NSV transmitted over channel state i has been held in the receiver buffer, another NSV sent over channel state j after one round-trip delay is received, and the convolutional decoding of the two NSV's fails. The average BER in the two received NSV's before convolutional decoding is then

$$\varepsilon(i,j) = (\varepsilon_i' + \varepsilon_j')/2, \tag{12}$$

where

$$\varepsilon_i' = \frac{\varepsilon_i}{P_{Ni}} \tag{13}$$

is the conditional BER given that the syndrome is nonzero, and

$$P_{Ni} = 1 - (1 - \varepsilon_i)^n - Q_i \tag{14}$$

is the probability of receiving a NSV sent while the channel was in state i. The BER at the output of the convolutional decoder is upper bounded by [24]

$$P_b(i,j) \le \frac{1}{2}\frac{\partial T(X,Y)}{\partial Y}\Big|X = 2\sqrt{\varepsilon(i,j)[1 - \varepsilon(i,j)]},\ Y = 1, \tag{15}$$

where $T(X, Y)$ is the generating function of the convolutional code.

The output of the convolutional decoder is then checked by the $(n-m,n-m-r)$ block code C_0. We assume that the errors at the output of the convolutional decoder are independent. The probability of undetected error after convolutional decoding in state ij is then upper bounded by

$$Q_{ij} \le 2^{-r}\{1 - 2[1 - P_b(i,j)]^{n-m} + [1 - 2P_b(i,j)]^{n-m}\}. \tag{16}$$

The probability of decoding failure is the probability of a nonzero syndrome in the block code C_0, and it is given by

$$P_{ij} = 1 - [1-P_b(i,j)]^{n-m} - Q_{ij}. \tag{17}$$

By arranging the states of Fig.3 in the order: 00, 01, 10, 11, e, c , the transition probabilities of the Markov chain are found in Appendix A and are summarized in the following transition probability matrix:

$$\mathbf{P} = \begin{bmatrix} p_{00,00} & p_{00,01} & 0 & 0 & p_{00,e} & p_{00,c} \\ 0 & 0 & p_{01,10} & p_{01,11} & p_{01,e} & p_{01,c} \\ p_{10,00} & p_{10,01} & 0 & 0 & p_{10,e} & p_{10,c} \\ 0 & 0 & p_{11,10} & p_{11,11} & p_{11,e} & p_{11,c} \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{18}$$

where, from (A.2), (A.3), and (A.4),

$$p_{00,00} = [1 - p_{01}(d)]\, P_{N0}\, P_{00}$$

$$p_{00,01} = p_{01}(d)\, P_{N1}\, P_{01}$$

$$p_{01,10} = [1 - p_{11}(d)]\, P_{N0}\, P_{10}$$

$$p_{01,11} = p_{11}(d)\, P_{N1}\, P_{11}$$

$$p_{10,00} = [1 - p_{01}(d)]\, P_{N0}\, P_{00}$$

$$p_{10,01} = p_{01}(d)\, P_{N1}\, P_{01}$$

$$p_{11,10} = [1 - p_{11}(d)]\, P_{N0}\, P_{10}$$

$$p_{11,11} = p_{11}(d)\, P_{N1}\, P_{11}$$

$$p_{00,e} = [1 - p_{01}(d)]\, [Q_0 + P_{N0}\, Q_{00}] + p_{01}(d)\, [Q_1 + P_{N1}\, Q_{01}] \tag{19}$$

$$p_{01,e} = [1 - p_{11}(d)]\, [Q_0 + P_{N0}\, Q_{10}] + p_{11}(d)\, [Q_1 + P_{N1}\, Q_{11}]$$

$$p_{10,e} = [1 - p_{01}(d)]\, [Q_0 + P_{N0}\, Q_{00}] + p_{01}(d)\, [Q_1 + P_{N1}\, Q_{01}]$$

$$p_{11,e} = [1 - p_{11}(d)]\, [Q_0 + P_{N0}\, Q_{10}] + p_{11}(d)\, [Q_1 + P_{N1}\, Q_{11}]$$

$$p_{00,c} = 1 - p_{00,00} - p_{00,01} - p_{00,e}$$

$$p_{01,c} = 1 - p_{01,10} - p_{01,11} - p_{01,e}$$

$$p_{10,c} = 1 - p_{10,00} - p_{10,01} - p_{10,e}$$

$$p_{11,c} = 1 - p_{11,10} - p_{11,11} - p_{11,e}\,,$$

$p_{01}(d)$ and $p_{11}(d)$ are the d-step channel transition probabilities given by (A.1), and d is the number of code vectors that can be transmitted during one channel round trip delay period.

Throughput efficiency calculation

Throughput efficiency is defined as the ratio of the average number of data vectors accepted by the receiver and delivered to the user per unit time to the total number of vectors that can be transmitted per unit time [26]. Let E[N] be the expected total number of transmissions (including the initial transmission and all retransmissions) required for a data vector to be successfully accepted by the receiver. Then the throughput of selective-repeat parity retransmission hybrid ARQ is given by [26]

$$\eta = \frac{1}{E[N]} \cdot \frac{n-m-r}{n}. \tag{20}$$

In practice, $m \ll n$ and $r \ll n$ and hence $(n-m-r)/n \simeq 1$. Let T be the expected number of retransmissions required for a data vector to be successfully decoded by the receiver. Then E[N] can be expressed as

$$E[N] = 1 + T = 1 + \sum_{i=0}^{1} T_i \, P_{Ni} \, p_i, \tag{21}$$

where T_i, $i = 0,1$, is the expected number of retransmissions required for a data vector to be successfully decoded by the receiver, given that the initial transmission was a NSV and was sent over channel state i, and $p_0 = 1 - p_1$ is the probability that the channel starts in state 0. Thus only T_i, $i = 0, 1$, must be determined to find η.

Referring to the Markov chain shown in Fig. 3, and assuming that an initial NSV has been received, we see that retransmissions will be needed to recover the information sequence associated with the initial NSV. As soon as an undetected error or error-free decoding occurs (i.e., the Markov chain reaches state e or state c), retransmissions will terminate, and the estimated information sequence will be delivered to the data sink. Therefore, the expected number of retransmissions is the *infinite-step chain mean absorption time* defined in Appendix B, and (B.9) can be applied to find T_i , $i = 0, 1$.

Suppose that the initial NSV was sent over channel state 0. To determine T_0, either state 00 or state 10 in Fig.3 can be used as our initial state, since both states assume that a NSV sent over channel state 0 is held in the receiver buffer and waiting to be processed upon reception of the first retransmission. Hence, the infinite-step chain mean absorption time (or equivalently, the expected number of retransmissions) starting from states 00 and 10 should be the same, i.e.,

$$T_0 = M_{00} = M_{10}. \tag{22.1}$$

Similarly,

$$T_1 = M_{01} = M_{11}. \tag{22.2}$$

Substituting the transition probabilities of (19) into (B.9), we obtain the following equations:

$$M_{00} = 1 + M_{00} \, p_{00,00} + M_{01} \, p_{00,01}$$

$$M_{01} = 1 + M_{10}\, p_{01,10} + M_{11}\, p_{01,11}.$$

Using (22.1) and (22.2), the above equations become

$$T_0\,(p_{00,00} - 1) + T_1\, p_{00,01} = -1 \tag{23.1}$$

$$T_0\, p_{01,10} + T_1\,(p_{01,11} - 1) = -1 \tag{23.2}$$

and their solutions are

$$T_0 = \frac{1 + p_{00,01} - p_{01,11}}{(1 - p_{00,00})\,(1 - p_{01,11}) - p_{00,01}\, p_{01,10}} \tag{24.1}$$

and

$$T_1 = \frac{1 + p_{01,10} - p_{00,00}}{(1 - p_{00,00})(1 - p_{01,11}) - p_{00,01}\, p_{01,10}}. \tag{24.2}$$

Therefore, from (21) and (24) we obtain

$$E[N] = 1 + \frac{[1 + p_{00,01} - p_{01,11}]\, P_{N0}\, p_0 + [1 + p_{01,10} - p_{00,00}]\, P_{N1}\, p_1}{(1 - p_{00,00})\,(1 - p_{01,11}) - p_{00,01}\, p_{01,10}}. \tag{25}$$

The throughput for selective-repeat parity retransmission hybrid ARQ can then be obtained by substituting (25) into (20).

For the special case when the channel is a BSC with BER ε , (25) can be simplified, after some calculations, to

$$E[N] = \frac{1 + P_N\,(1 - P_f)}{1 - P_N\, P_f}, \tag{26}$$

where

$$P_N \simeq 1 - (1 - \varepsilon)^n \tag{27}$$

is the probability of receiving a NSV and

$$P_f \simeq 1 - (1 - p_b)^{n-m} \tag{28}$$

is the probability of a convolutional decoding failure on two NSVs,

$$p_b \le \frac{1}{2}\frac{\partial T(X,Y)}{\partial Y}\Big|_{X=2\sqrt{\varepsilon'(1-\varepsilon')},\ Y=1}$$

and

$$\varepsilon' \simeq \frac{\varepsilon}{1-(1-\varepsilon)^n}.$$

The throughput on a *BSC* is then

$$\eta = \frac{1 - P_N P_f}{1 + P_N (1 - P_f)} \frac{n-m-r}{n}. \tag{29}$$

In [17], only a lower bound was obtained for a parity retransmission hybrid ARQ scheme on a BSC.

Undetected error probability calculation

The *undetected error probability* is the average probability of decoding error, given that the receiver accepts a data vector. We denote this average probability by P_{ud}. It is easy to see that

$$P_{ud} = \sum_{i=0}^{1} \{Q_i + U_i P_{Ni}\} P_i, \tag{30}$$

where U_i, $i = 0,1$, is the probability of undetected error in the data vector accepted by the receiver, given that the initial transmission was a NSV and sent over channel state i.

Referring to Fig. 3, given that an initial NSV was received, undetected errors will occur if and only if the Markov chain enters state e. Therefore, finding U_i is equivalent to finding the *infinite-step absorption probability* of the Markov chain (see Appendix B). Let $A_{00,e}$, $A_{01,e}$, $A_{10,e}$, and $A_{11,e}$ be the probabilities that the chain of Fig.3, starting at states 00, 01, 10, and 11, respectively, will eventually be absorbed by state e. By an argument similar to the throughput calculation, we can show that

$$U_0 = A_{00,e} = A_{10,e} \tag{31.1}$$

$$U_1 = A_{01,e} = A_{11,e}. \tag{31.2}$$

Substituting the transition probabilities of (19) into (B.8), it follows that

$$A_{00,e} = A_{00,e}\, p_{00,00} + A_{01,e}\, p_{00,01} + p_{00,e}$$

$$A_{01,e} = A_{10,e}\, p_{01,10} + A_{11,e}\, p_{01,11} + p_{01,e}.$$

From (31), the above equations reduce to

$$U_0 (1 - p_{00,00}) - U_1 p_{00,01} = p_{00,e} \tag{32.1}$$

$$-U_0 p_{01,10} + U_1(1 - p_{01,11}) = p_{01,e}. \tag{32.2}$$

Solving (32) we obtain

$$U_0 = \frac{p_{00,e}(1-p_{01,11})+p_{01,e}p_{00,01}}{(1-p_{00,00})(1-p_{01,11})-p_{00,01}p_{01,10}} \tag{33.1}$$

$$U_1 = \frac{p_{01,e}(1-p_{00,00})+p_{00,e}p_{01,10}}{(1-p_{00,00})(1-p_{01,11})-p_{00,01}p_{01,10}}. \tag{33.2}$$

Thus, from (30) and (33), the probability of undetected error for selective-repeat parity retransmission hybrid ARQ with an infinite receiver buffer is given by

$$P_{ud} = \left[Q_0+\frac{p_{00,e}(1-p_{01,11})+p_{01,e}p_{00,01}}{(1-p_{00,00})(1-p_{01,11})-p_{00,01}p_{01,10}}P_{N0}\right]p_0 + \left[Q_1+\frac{p_{01,e}(1-p_{00,00})+p_{00,e}p_{01,10}}{(1-p_{00,00})(1-p_{01,11})-p_{00,01}p_{01,10}}P_{N1}\right]p_1. \tag{34}$$

If the channel is a BSC with BER ε, (34) can be simplified to

$$P_{ud} = Q + \frac{Q + P_N P_e}{1 - P_N P_f}, \tag{35}$$

where P_N and P_f are given by (27) and (28), respectively,

$$Q \le 2^{-(m+r)}[1 - 2\,(1 - \varepsilon)^n + (1 - 2\,\varepsilon)^n] \tag{36}$$

is the probability of receiving a ZSV with an undetected error pattern, and

$$P_e \le 2^{-r}[1 - 2\,(1 - p_b)^{n-m} + (1 - 2p_b)^{n-m}] \tag{37}$$

is the probability of undetected error after convolutional decoding of two NSVs.

5. System Throughput and Undetected Error Probability with a Finite Receiver Buffer

In this section we analyze the throughput and the probability of undetected error for the selective-repeat parity retransmission hybrid ARQ with a finite receiver buffer. Let B be the number of code vectors that the receiver buffer can store. The performance is analyzed for the case when

$$B = l \cdot d, \qquad l = 1,2,3,.... \tag{38}$$

The system with receiver buffer size B operates as follows. Normally, the transmitter sends code vectors continuously to the receiver. The receiver checks the syndrome of each received code vector. If the syndrome is zero, the received vector is assumed to be error-free and is delivered to the user, and an ACK signal is sent to the transmitter. When the channel is quiet, data transmission proceeds smoothly; error-free vectors are delivered to the user in consecutive order and the receiver buffer is empty. The receiver is said to be in the *normal phase* if the receiver buffer is empty.

When a received code vector is detected in error (NSV) while the receiver is in the normal phase, the receiver enters the *blocked phase* and sends a NACK to the transmitter. The NSV is then stored in the receiver buffer for error correction at a later time. In the blocked phase, the receiver continues to check the syndrome of each incoming received code vector, sends an ACK to the transmitter for each received ZSV, and sends a NACK to the transmitter for each received NSV. The received vectors, no matter whether they are ZSVs or NSVs, are stored in the receiver buffer until they are ready to be released to the data sink. In the blocked phase, no vector is delivered to the data sink until the earliest received NSV is accepted by the data sink.

If the earliest NSV is recovered within l retransmissions, the receiver then starts to deliver this vector and the subsequent ZSVs (which are held in the receiver buffer) to the data sink in order until the next NSV is encountered. This vector then becomes the earliest NSV. If all the vectors held in the receiver buffer are released to the data sink after the earliest NSV has been recovered, the receiver buffer becomes empty again and the receiver returns to the normal phase.

If the receiver fails to recover the earliest NSV after l retransmissions, no further retransmissions are allowed. The receiver simply delivers the erroneous decoded vector to the data sink and sends an ACK to the transmitter. Because the receiver has a buffer size $B = l \cdot d$, buffer overflow will never occur, and the system operates in the same way as though the receiver had an infinite buffer size, except for the *forced* vector delivery upon receiving the *lth* retransmission. The tradeoff in system performance by limiting the number of retransmissions to l will be an increased throughput and a decreased system reliability compared with systems without this limitation.

Throughput efficiency calculation

Let $E[N(l)]$ be the expected number of transmissions needed to deliver a code vector to the user within $l+1$ transmissions (including the initial transmission and up to l retransmissions). The throughput of the selective-repeat parity retransmission hybrid ARQ with receiver buffer size $B = l \cdot d$ is then

$$\eta = \frac{1}{E[N(l)]} \frac{n-m-r}{n}. \tag{39}$$

By a similar argument as in the infinite receiver buffer case, we obtain

$$E[N(l)] = 1 + T_0(l)\, P_{N0}\, p_0 + T_1(l)\, P_{N1}\, p_1 \,, \tag{40}$$

where $T_i(l)$, $i = 0, 1$, is the expected number of retransmissions required for a data vector to be successfully decoded within l retransmissions, given that the initial vector was a NSV and sent over channel state i.

We readily recognize that finding $T_i(l)$ is equivalent to finding the Markov chain mean absorption time within l transitions (see Appendix B), so that (B.4) and (B.6) can be used. Note that the computation can be reduced by combining state e and state c in Fig.3 into a single state, called state s, since both states e and c are absorbing states. Thus state s is simply the state corresponding successful decoding. The new transition probability matrix, obtained by merging states e and c into state s, is given by (see (18))

$$\mathbf{P}' = \begin{bmatrix} p_{00,00} & p_{00,01} & 0 & 0 & p_{00,s} \\ 0 & 0 & p_{01,10} & p_{01,11} & p_{01,s} \\ p_{10,00} & p_{10,01} & 0 & 0 & p_{10,s} \\ 0 & 0 & p_{11,10} & p_{11,11} & p_{11,s} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{41.1}$$

where

$$p_{ij,s} = p_{ij,e} + p_{ij,c}. \tag{41.2}$$

Let $M_{00}(l)$ and $M_{01}(l)$ be the mean absorption time within l transitions of the Markov chain described by (41), conditioned on the chain starting in state 00 and state 01, respectively. From (B.4) we have

$$T_0(l) = M_{00}(l) = p_{00,s} + \sum_{n=2}^{l} n\, [p_{00,s}(n) - p_{00,s}(n-1)\,] \,, \tag{42.1}$$

$$T_1(l) = M_{01}(l) = p_{01,s} + \sum_{n=2}^{l} n\,[p_{01,s}(n) - p_{01,s}(n-1)\,], \tag{42.2}$$

where $p_{00,s}(n)$ and $p_{01,s}(n)$ are the n-step transition probabilities from state 00 and state 01, respectively, to state s, and they can be determined from (41), (19), and (10). The system throughput of selective-repeat parity retransmission hybrid ARQ with receiver buffer size $B = l \cdot d$, from (39), (40), and (42), is given by

$$\eta = \frac{(n-m-r)/n}{1 + \sum_{i=0}^{1}\left\{p_{0i,s} + \sum_{n=2}^{l} n\,[p_{0i,s}(n) - p_{0i,s}(n-1)]\right\} P_{Ni}\,p_i}. \tag{43}$$

Undetected error probability calculation

The probability of undetected error for selective-repeat parity retransmission hybrid ARQ with an infinite receiver buffer is given in (30). The probability of undetected error for the system with receiver buffer size $B = l \cdot d$ can be derived from (30) with a slight modification. It is given by

$$P_{ud} = \sum_{i=0}^{1}\left\{Q_i + [U_i(l) + E_i(l)]\,P_{Ni}\right\} p_i, \tag{44}$$

where $U_i(l)$ is the probability of undetected error within l retransmissions, given that the initial transmission was a NSV and sent over channel state i. $E_i(l)$ is the probability of unsuccessful decoding on the *lth* retransmission, given that the initial transmission was a NSV and sent over channel state i. Since only up to l retransmissions are permitted in the system, if the *lth* retransmission results in a decoding failure, the erroneously decoded data vector will be delivered to the data sink. Hence, it is reasonable to regard this decoding failure as an undetected error, and its probability , $E_i(l)$, should be included in (44).

Let $A_{00,e}(l)$ and $A_{01,e}(l)$ be the probabilities that the Markov chain, starting at states 00 and 01, respectively, will be absorbed by state e within l transitions. Clearly, $U_0(l) = A_{00,e}(l)$ and $U_1(l) = A_{01,e}(l)$. It follows from (B.2) that

$$U_0(l) = A_{00,e}(l) = p_{00,e}(l) \tag{45.1}$$

and

$$U_1(l) = A_{01,e}(l) = p_{01,e}(l), \tag{45.2}$$

where $p_{0i,e}(l)$, $i = 0, 1$, is the l-step transition probability from state $0i$ to state e. We also have

$$E_i(l) = Pr\left\{\text{The chain reaches 00, 01, 10, or 11 after } l \text{ steps} \mid \text{The chain starts in } 0i\right\}$$

$$= p_{0i,00}(l) + p_{0i,01}(l) + p_{0i,10}(l) + p_{0i,11}(l). \tag{46}$$

Combining (45) and (46), we obtain

$$U_i(l) + E_i(l) = 1 - p_{0i,c}(l), \tag{47}$$

which is the probability that the chain starting in state $0i$, $i = 0, 1$, will not reach state c within l transitions. In other words, $U_i(l) + E_i(l)$ is simply the probability that the initial NSV will not be decoded correctly within l retransmissions. Thus, the probability of undetected error for selective-repeat parity retransmission hybrid ARQ with receiver buffer size $B = l \cdot d$, from (44) and (47), is given by

$$P_{ud} = \sum_{i=0}^{1} \left\{Q_i + [1 - p_{0i,c}(l)]\, P_{Ni}\right\} p_i. \tag{48}$$

6. Examples

In this section, we plot the performance of selective-repeat parity retransmission hybrid ARQ for the following parameters:

C_0 : (1024, 1000) binary code

C_1 : (2, 1, 6) $d_f = 10$ convolutional code

Channel round trip delay : d = 128.

Four curves, numbered 0 through 3, appear in each plot. They correspond to four values of the burst duty cycle, which was defined in section 3 as the parameter which determines the burstiness of the channel:

Curve 0 : $p_1 = 1$ (Stationary channel)

Curve 1 : $p_1 = 0.25$ (Diffuse burst channel)

Curve 2 : $p_1 = 0.1$

Curve 3 : $p_1 = 0.05$ (Dense burst channel) .

In the performance calculation, we found that, for a given p_1 and $\bar{\varepsilon}$, different values of $\bar{b}$ have a very small influence on the throughput and undetected error probability. Therefore, we considered only the case when $p = p' = p_1$, i.e., a BI channel. Figs. 4 and 5 show the throughput and the probability of undetected error for selective-repeat parity retransmission hybrid ARQ with infinite receiver buffer. Fig. 6 shows the probalility of undetected error with receiver buffer size $B = 5d$. The system throughput is essentially the same as in the infinite receiver buffer case.

From Fig. 4 we observe that the system throughput is much better for a given average BER if the errors occur in bursts, since the errors are then concentrated in fewer vectors and the convolutional code is still powerful enough to decode them. From Fig. 5 we see that the burstiness of the channel has a limited influence on the undetected error probalility, except in the range of average BERs $4\times10^{-4} \le \bar{\varepsilon} \le 10^{-2}$. In this range, and for a bursty channel, most of the errors are concentrated in the vectors received while the channel is noisy. The convolutional code is powerful enough to decode most of these noisy vectors reliably. On the other hand, if the channel is stationary, the vectors which are received during the quiet state contain more errors, and there are more undetected errors than in the non-stationary case, because of the limited power of the block decoder.

Fig. 6 shows the degradation in system reliability when the receiver buffer size is limited to five times the channel round trip delay ($B = l \cdot d = 5d$). The degradation becomes obvious for average BERs in the range $\bar{\varepsilon} \ge 4\times10^{-3}$. For small l and high $\bar{\varepsilon}$, although most errors can be corrected within l retransmissions, a small percentage of error vectors cannot be corrected, and those error vectors will have a serious effect on the probability of undetected error. Obviously, as l becomes large, the system becomes more and more reliable.

7. Discussion and Conclusions

Wang and Lin proposed a hybrid ARQ scheme with parity retransmission using two block codes C_0 and C_1 [17]. C_0 is an (n,k) high-rate error-detecting code and C_1 is a half-rate $(2n,n)$ code which is designed for error correction only. However, C_1 must be invertible and an inverse operation is required at the decoder to recover the

information data vector from the retransmitted parity check vector. In [17], the performance of hybrid ARQ schemes using both a rate 1/2 invertible block code and a rate 1/2 convolutional code were analyzed for a BSC. The results indicate that convolutional codes provide a higher throughput than block codes, especially on very noisy channels (BERs around 10^{-2}). However, since these results were obtained for a BSC, and since parity retransmission hybrid ARQ schemes are designed for use on non-stationary channels, a more detailed analysis was needed. In this paper, we have presented a thorough analysis of a parity retransmission hybrid ARQ scheme using convolutional codes for a non-stationary channel, assuming both infinite and finite receiver buffers. Results show that high throughput efficiencies and low undetected error probabilities can be maintained over a wide range of channel parameters and that the throughput efficiency improves as the channel becomes burstier in nature. In addition, the undetected error probability performance is better for burstier channels. These results substantiate the claim that parity retransmission hybrid ARQ schemes using convolutional codes are particularly well suited for use on non-stationary channels.

Unlike block codes, "invertibility" is not required for convolutional codes in a hybrid ARQ scheme with parity retransmission. Therefore, convolutional codes with a variety of code rates can be incorporated into hybrid ARQ schemes. Hybrid ARQ schemes using rate 2/3 and 3/4 convolutional codes have been proposed and analyzed over non-stationary channels in [20]. The rate 2/3 and 3/4 schemes provide much better throughput than the rate 1/2 schemes at the cost of a more complex Viterbi decoder, a larger buffer, and a more complex buffer strategy. For instance, assuming the duty cycle of the noisy burst is $p_1 = 0.05$, the average channel BER is $\bar{\varepsilon} = 10^{-2}$, and the memory order of the convolutional code is 3, the throughput is equal to 0.6 for the rate 1/2 scheme, 0.74 for the rate 2/3 scheme, and 0.8 for the rate 3/4 scheme [20]. These higher rate hybrid ARQ schemes are very attractive for use on high speed non-stationary channels, such as satellite communication channels.

Appendix A

In this appendix we derive the transition probabilities of the Markov chain shown in Fig.3.

Transition probability from state *kj* to state *ji*, $p_{kj,ji}$

Let d be the number of code vectors that can be transmitted during one channel round trip delay period. Since the retransmitted vector is received after one channel round trip delay , in the following calculations we need to know the channel d-step transition probabilities. The channel d-step transition probabilities $p_{01}(d)$ and $p_{11}(d)$ of being in state 1 d time frames after being in state 0 and state 1, respectively, are given by [25]

$$p_{01}(d) = \frac{p}{1+p-p'} - \frac{p}{1+p-p'}\,(p'-p)^d, \tag{A.1.1}$$

and

$$p_{11}(d) = \frac{p}{1+p-p'} + \frac{(1-p')}{1+p-p'}\,(p'-p)^d. \tag{A.1.2}$$

Let us consider the transition from state 01 to state 10. State 01 means that two NSVs $\hat{V}_1(x)$ and $\hat{V}_2(x)$, sent over channel states 0 and 1, respectively, have been received and that decoding has failed. $\hat{V}_1(x)$ is then discarded and $\hat{V}_2(x)$ is held in the receiver buffer. The transition from state 01 to state 10 means that a new NSV $\hat{V}_3(x)$, sent over channel state 0 after a round trip delay, has been received with probability $[\,1 - p_{11}(d)\,]\,P_{N0}$, and that convolutional decoding of the two NSVs, $\hat{V}_2(x)$ and $\hat{V}_3(x)$, has failed with probability P_{10}. The transition probability from state 01 to state 10 is therefore given by

$$p_{01,10} = [1 - p_{11}(d)]\,P_{N0}\,P_{10}.$$

By a similar argument we obtain

$$p_{00,00} = [1 - p_{01}(d)]\,P_{N0}\,P_{00}$$

$$p_{00,01} = p_{01}(d)\,P_{N1}\,P_{01}$$

$$p_{01,10} = [1 - p_{11}(d)]\,P_{N0}\,P_{10}$$

$$p_{01,11} = p_{11}(d)\,P_{N1}\,P_{11}$$

$$p_{10,00} = [1 - p_{01}(d)]\, P_{N0}\, P_{00}$$

$$p_{10,01} = p_{01}(d)\, P_{N1}\, P_{01} \qquad \text{(A.2)}$$

$$p_{11,10} = [1 - p_{11}(d)]\, P_{N0}\, P_{10}$$

$$p_{11,11} = p_{11}(d)\, P_{N1}\, P_{11}.$$

Transition probability from state *ij* to state *e*, $p_{ij,e}$

Consider the transition from state 01 to state *e*. In state 01, the decoding of the two NSVs $\hat{V}_1(x)$ and $\hat{V}_2(x)$ sent over channel states 0 and 1, respectively, has failed. $\hat{V}_1(x)$ has been discarded and replaced by a retransmitted vector $\hat{V}_3(x)$, which was received one round trip delay after the reception of $\hat{V}_2(x)$. The transition from state 01 to state *e* means an undetected error is made in the syndrome check of $\hat{V}_3(x)$, or if $\hat{V}_3(x)$ is a NSV, in the decoding of $\hat{V}_2(x)$ and $\hat{V}_3(x)$. Note that $\hat{V}_3(x)$ may be sent over channel state 0 or 1. Consider the case when the channel is in state 0. The probability of such an event is the *d*-step transition probability in the channel model from state 1 to state 0, [1 - $p_{11}(d)$]. An undetected error is made when performing the syndrome check on $\hat{V}_3(x)$ with probability Q_0. If the syndrome check on $\hat{V}_3(x)$ is nonzero (with probability P_{N0}), another undetected error can be made in the convolutional decoding of $\hat{V}_2(x)$ and $\hat{V}_3(x)$ with probability Q_{10}. The undetected error probability, given that $\hat{V}_3(x)$ is received while the channel is in state 0, is therefore

$$E_0 = Q_0 + P_{N0}\, Q_{10}.$$

Similarly, the probability of making an undetected error, given that $\hat{V}_3(x)$ is received when the channel is in state 1, is

$$E_1 = Q_1 + P_{N1}\, Q_{11}.$$

The transition probability from state 01 to state *e* is then obtained by averaging E_0 and E_1 :

$$p_{01,e} = [1 - p_{11}(d)]\, E_0 + p_{11}(d)\, E_1.$$

The transition probabilities from the three other states to state *e* are obtained similarly, and we have

$$p_{00,e} = [1 - p_{01}(d)]\, (Q_0 + P_{N0}\, Q_{00}) + p_{01}(d)\, (Q_1 + P_{N1}\, Q_{01})$$

$$p_{01,e} = [1 - p_{11}(d)]\ (Q_0 + P_{N0}\ Q_{10}) + p_{11}(d)\ (Q_1 + P_{N1}\ Q_{11})$$

$$p_{10,e} = [1 - p_{01}(d)]\ (Q_0 + P_{N0}\ Q_{00}) + p_{01}(d)\ (Q_1 + P_{N1}\ Q_{01}) \qquad \text{(A.3)}$$

$$p_{11,e} = [1 - p_{11}(d)]\ (Q_0 + P_{N0}\ Q_{10}) + p_{11}(d)\ (Q_1 + P_{N1}\ Q_{11}).$$

Transition probability from state *ij* to state *c*, $p_{ij,c}$

Realizing that the transitions from state *ij* to states *jk*, *e*, and *c* are mutually exclusive and collectively exhaustive, the sum of the corresponding transition probabilities must add up to 1. Therefore

$$p_{00,c} = 1 - p_{00,00} - p_{00,01} - p_{00,e}$$

$$p_{01,c} = 1 - p_{01,10} - p_{01,11} - p_{01,e}$$

$$p_{10,c} = 1 - p_{10,00} - p_{10,01} - p_{10,e} \qquad \text{(A.4)}$$

$$p_{11,c} = 1 - p_{11,10} - p_{11,11} - p_{11,e}.$$

Appendix B

In this appendix we present some results concerning Markov chains which are used throughout the paper.

First passage probability

The *n-step first passage probability*, denoted by $f_{j,b}(n)$, is defined as the probability that a Markov chain starting from state j will be in state b for the first time after n transitions. If b is an absorbing state, these probabilities can be found directly from the n-step transition probabilities as follows:

$$\begin{aligned} f_{j,b}(1) &= p_{j,b}(1) = p_{j,b} \\ f_{j,b}(2) &= p_{j,b}(2) - p_{j,b}(1) \\ f_{j,b}(3) &= p_{j,b}(3) - p_{j,b}(2) \qquad \text{(B.1)} \\ &\cdot \\ &\cdot \\ &\cdot \\ f_{j,b}(n) &= p_{j,b}(n) - p_{j,b}(n-1). \end{aligned}$$

To prove (B.1), we observe that the one-step first passage probability is the same as the one-step transition probability. For $n = 2$, the two-step transition probability $p_{j,b}(2)$ contains the probability of visiting state b immediately after the first transition and remaining in state b during the second transition. Hence, the probability of this event, $p_{j,b}(1) \cdot p_{b,b}(1)$, should be subtracted from $p_{j,b}(2)$ to obtain the two-step first passage probability, i.e., $f_{j,b}(2) = p_{j,b}(2) - p_{j,b}(1) \cdot p_{b,b}(1)$. Since b is an absorbing state (i.e., $p_{b,b}(1) = 1$), we have $f_{j,b}(2) = p_{j,b}(2) - p_{j,b}(1)$. The rest of (B.1) is based on similar reasoning.

Once the first passage probabilities have been determined, the following important quantities can be evaluated.

Absorption probability and mean absorption time

The *l-step absorption probability,* denoted by $A_{j,b}(l)$, is the probability that a Markov chain initially in state j will be absorbed by absorbing state b *within l transitions.* It follows from (B.1) that

$$A_{j,b}(l) = \sum_{n=1}^{l} f_{j,b}(n) = p_{j,b}(l). \tag{B.2}$$

The *infinite-step absorption probability,* denoted by $A_{j,b}$, is the probability that a chain starting in state j will *eventually* be absorbed by absorbing state b. Thus from (B.2),

$$A_{j,b} = A_{j,b}(\infty) = p_{j,b}(\infty). \tag{B.3}$$

Note that if the Markov chain contains only one absorbing state, $A_{j,b} = 1$, since the chain will eventually reach b and be trapped there.

We denote the *l-step mean absorption time* by $M_{j,b}(l)$. It is defined as the mean time that a Markov chain starting in state j will be absorbed by absorbing state b *within l transitions.* From (B.1) we obtain

$$M_{j,b}(l) = \sum_{n=1}^{l} n\, f_{j,b}(n) = p_{j,b} + \sum_{n=2}^{l} n\, [p_{j,b}(n) - p_{j,b}(n-1)]. \tag{B.4}$$

The *infinite-step mean absorption time,* denoted $M_{j,b}$, is then

$$M_{j,b} = M_{j,b}(\infty) = p_{j,b} + \sum_{n=2}^{\infty} n\, [p_{j,b}(n) - p_{j,b}(n-1)]. \tag{B.5}$$

$M_{j,b}$ is the mean period of time after the chain leaves state j until it is eventually absorbed by state b.

Let S be the set of absorbing states in a Markov chain. Define the *l-step chain mean absorption time* as

$$M_j(l) = \sum_{b \in S} M_{j,b}(l), \tag{B.6}$$

and the *infinite-step chain mean absorption time* as

$$M_j = \sum_{b \in S} M_{j,b}. \tag{B.7}$$

Obviously, $M_j(l)$ is the mean time of absorption within l-transitions if the chain starts from a given initial state j, and M_j is the mean time required for a Markov chain to be

eventually absorbed if it starts in state j.

Easy ways of finding $A_{j,b}$ and M_j

Because the calculations of $A_{j,b}$ and M_j involve the infinite-step transition probability, $p_{j,b}(\infty)$, direct evaluation of $A_{j,b}$ and M_j from (B.3) and (B.7) become impractical. However, they can be determined in alternative ways. Let N be the number of states in the chain. The $A_{j,b}$'s are related by the following set of equations [27]:

$$A_{j,b} = \sum_{i=1}^{N} A_{i,b}\, p_{j,i} \qquad \text{for } all\ j \in S. \tag{B.8}$$

Note that, when the initial state is b, it is already absorbed in b, and hence $A_{b,b} = 1$, whereas when the initial state is some other absorbing state, say a, then it will never be absorbed by b, and hence $A_{a,b} = 0$. For all other initial states, a set of simultaneous equations may be written from (B.8) whose solution is $A_{j,b}$.

The infinite-step chain mean absorption times, M_j, can be determined by solving the following set of equations [27]:

$$M_j = 1 + \sum_{i=1}^{N} M_i\, p_{j,i} \qquad \text{for } all\ j \in S. \tag{B.9}$$

Observe that if state i is an absorbing state, the chain is already in an absorbing state, and $M_i = 0$.

REFERENCES

[1] I. M. Jacobs, R. Binder, and E. V. Hoversten, "General Purpose Packet Satellite Networks ", *Proc. IEEE,* VOL. 66, pp. 1448- 1448, Nov. 1978.

[2] R. E. Kahn , S. A. Gronemeyer, J. Burchfiel, and R. E. Kunzelman, "Advances in Packet Radio Technology", *Proc. IEEE,* VOL. 66, pp. 1468-1497, Nov. 1978.

[3] S. Lin and P. S. Yu, "A Hybrid ARQ Scheme with Parity Retransmission for Error Control of Satellite Channels", *IEEE Trans. Comm.,*VOL. COM-30, pp.1701-1719, July 1982.

[4] K. Brayer, "Error Control Techniques Using Binary Symbol Burst Codes," *IEEE Trans. Comm.*, COM-16, pp. 199-214, April 1968.

[5] E. Y. Rocher and R. L. Pickholtz, "An Analysis of Effectiveness of Hybrid Transmission Schemes", *IBM J. Res. Dev., pp. 426-433, July 1970.*

[6] A. R. K. Sastry, "Performance of Hybrid Error Control Schemes on Satellite Channels", *IEEE Trans. Comm.,* COM-23, pp. 689-694, July 1975.

[7] A. R. K. Sastry and L. N. Kanal, "Hybrid Error Control Using Retransmission and Generalized Burst-Trapping Codes", *IEEE Trans. Comm.,* COM-24, pp. 385-393, April 1976.

[8] D. Haccoun, J. Conan, and G. Golly, "Node to Node Protocols on a High Speed Full Duplex Satellite Link", *NTC'78 Conf. Rec., Dec. 1978.*

[9] C. S. K. Leung and A. Lam, "Forward Error Correction for an ARQ Scheme", *IEEE Trans. Comm.,* COM-29, pp. 1514-1519, Oct. 1981.

[10] K. E. Perry and J. M. Wozencraft, "SECO: A Self-Regulating Error-Correcting Coder-Decoder", *IRE Trans. Inform. Th.,* IT-8, pp. 128-135, Sept. 1962.

[11] H. Yamamoto and K. Itoh, "Viterbi Decoding Algorithm for Convolutional Codes with Repeat Request", *IEEE Trans. Inform. Th.,* IT-26, pp. 540-547, Sept. 1980.

[12] A. Drukarev and D. J. Costello, Jr., "Hybrid ARQ Error Control Using Sequential Decoding ", *IEEE Trans. Inform. Th.,* IT-29, pp. 521-535, July 1983.

[13] D. M. Mandelbaum, "Adaptive Feedback Coding Scheme Using Incremental Redundancy", *IEEE Trans. Inform. Th.,* IT-20, pp. 388-389, May 1974.

[14] R. A. Comroe and D. J. Costello, Jr., "ARQ Schemes for Data Transmission in Mobile Radio Systems", *IEEE J. Selected Areas Commun.*, SAC-2, pp. 472-481, July 1984.

[15] J. J. Metzner, "Improvements in Block Retransmission Schemes", *IEEE Trans. Comm.*, COM-27, pp. 524-532, Feb. 1979.

[16] T. C. Ancheta, "Convolutional Parity Check Automatic Repeat Request", *IEEE Int. Symp. on Inform. Th.*, Grignano, Italy, June 1979.

[17] Y. M. Wang and S. Lin, "A Modified Selective-Repeat Type-II Hybrid ARQ System and Its Performance Analysis", *IEEE Trans. Comm.*, COM-31, pp.593-608, May 1983.

[18] M. J. Miller and S. Lin, "The Analysis of Some Selective-Repeat ARQ Schemes with Finite Receiver Buffer", *IEEE Trans. Comm.*, COM-29, pp.1307-1315, Sept. 1981.

[19] E. N. Gilbert, "Capacity of Burst Noise Channels", *Bell Sys. Tech. J.*, VOL-39, pp. 1253-1256, 1960.

[20] L. R. Lugand, "ARQ Schemes Using Convolutional Codes and Viterbi Decoding Over Non-Stationary Channels", Ph.D. Dissertation, Department of Electrical Engineering, Illinois Institute of Tech., Chicago, IL, Dec. 1982.

[21] J. L. Massey, *Coding Techniques for Digital Communications,* notes for a tutorial session, 1973 IEEE Int. Conf. Comm., Seattle, Washington, June 1973.

[22] R. J. McEliece and W. E. Stark, "Channels with Block Interference", *IEEE Trans. Inform. Th.*, IT-30, pp. 44-53, Jan. 1984.

[23] T. Kasami, T. Klove, and S. Lin, "On the Probability of Undetected Error of Linear Block Codes", *Globecom'82 Conf. Rec.*, pp.E7.4.1-E.7.4.3, Miami, Florida, Dec. 1982.

[24] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding,* New York: McGraw-Hill, 1979.

[25] H. Larson and B. Shubert, *Probabilistic Models in Engineering and Sciences,* VOL.II, John Wiley and Sons, 1979.

[26] S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications,* Prentice Hall, Inc., 1982.

[27] A. H-S. Ang and W. H. Tang, *Probability Concepts in Engineering Planning and Design, VOL.II- Decision, Risk, and Reliability,* John Wiley and Sons, New York, 1984.

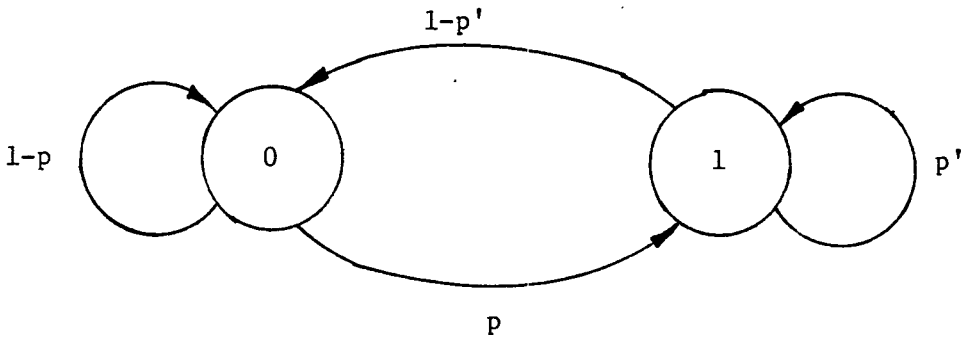


Fig. 1 A two-state Markov chain non-stationary channel model.

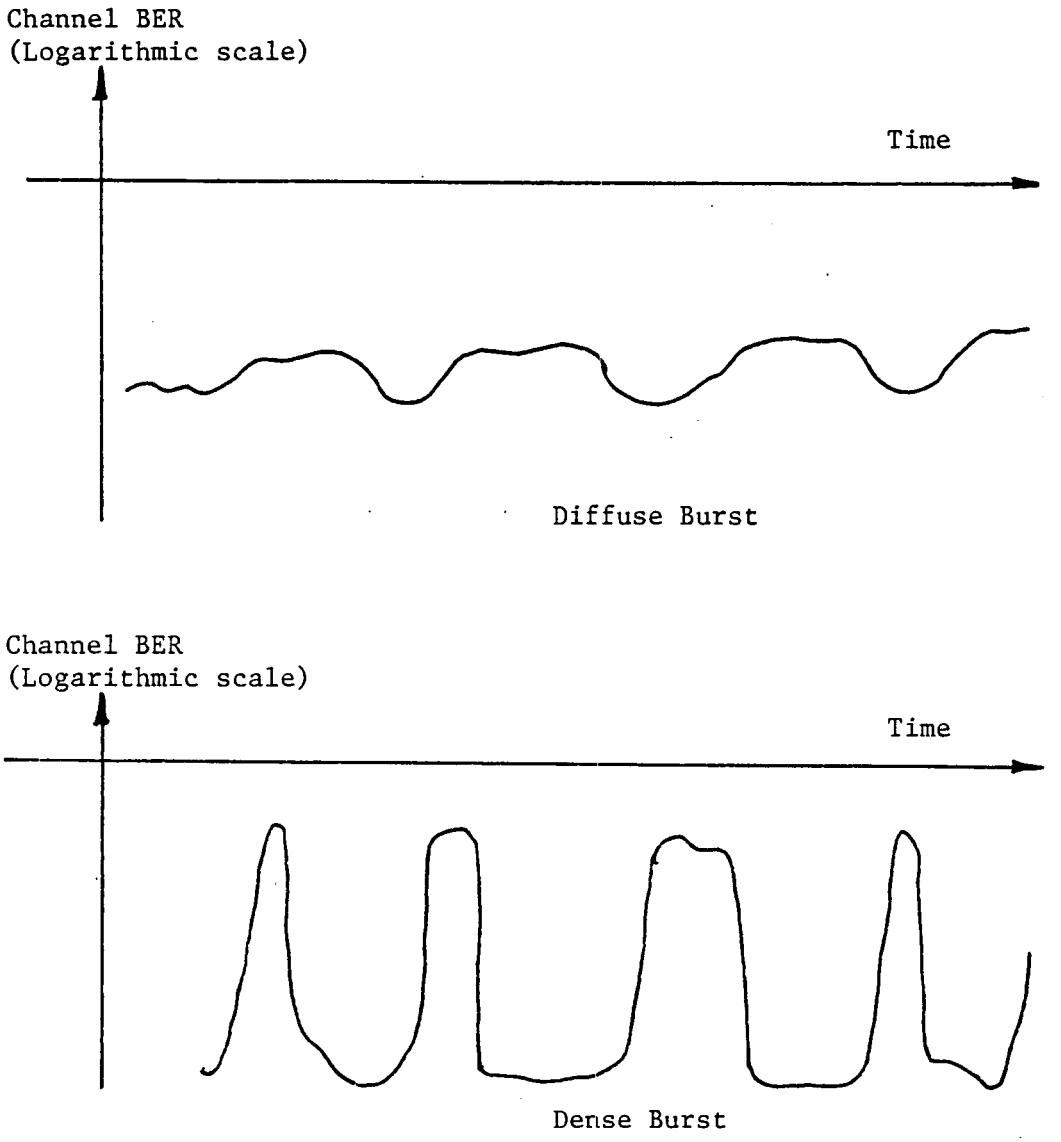


Figure 2 Diffuse and Dense Burst Channel Models

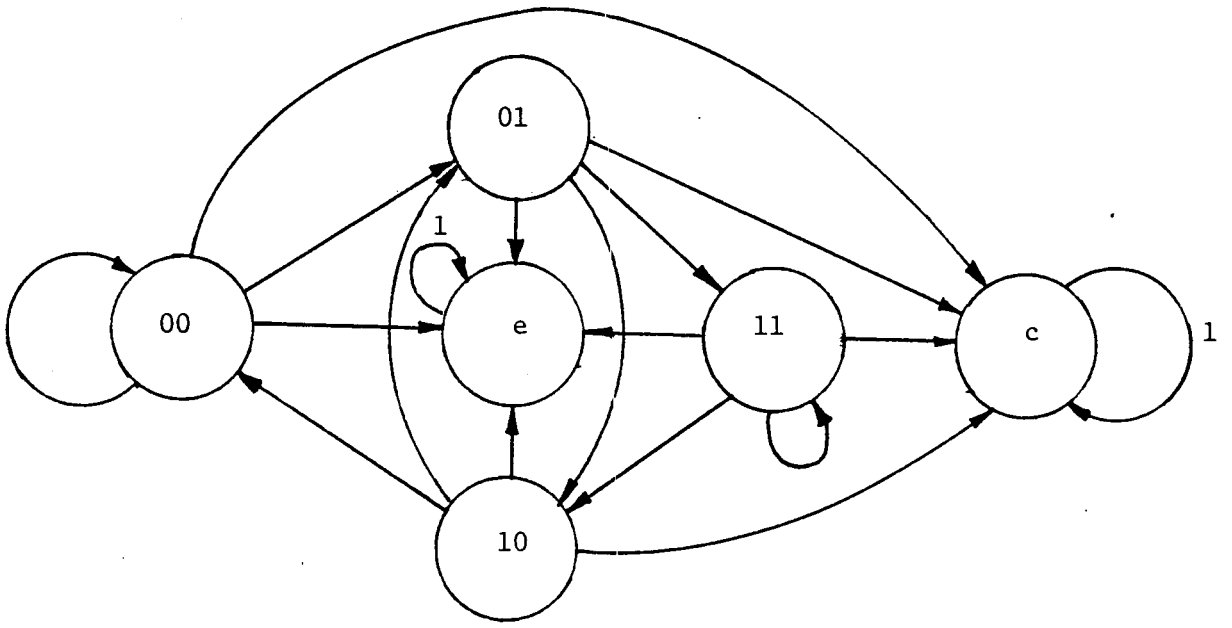


Fig. 3 Receiver's decoding status after receiving an initial NSV.

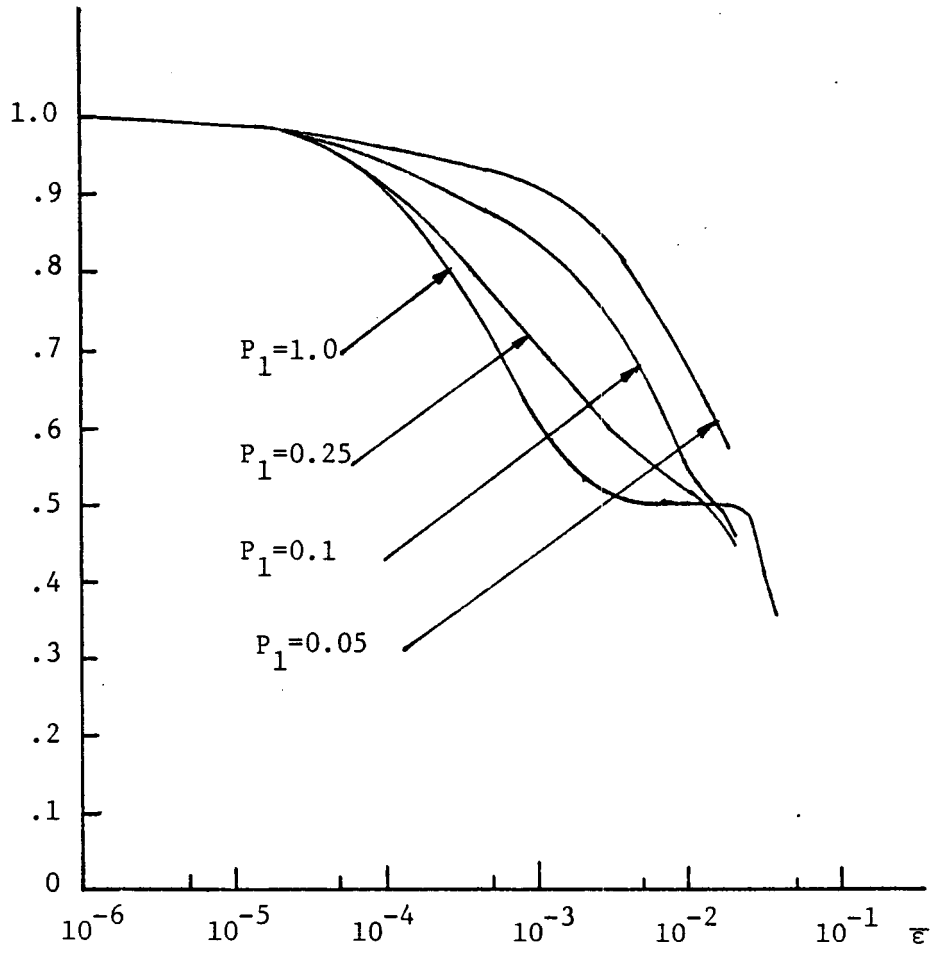


Fig. 4 Throughput performance of rate 1/2 selective-repeat ARQ with infinite receiver buffer.

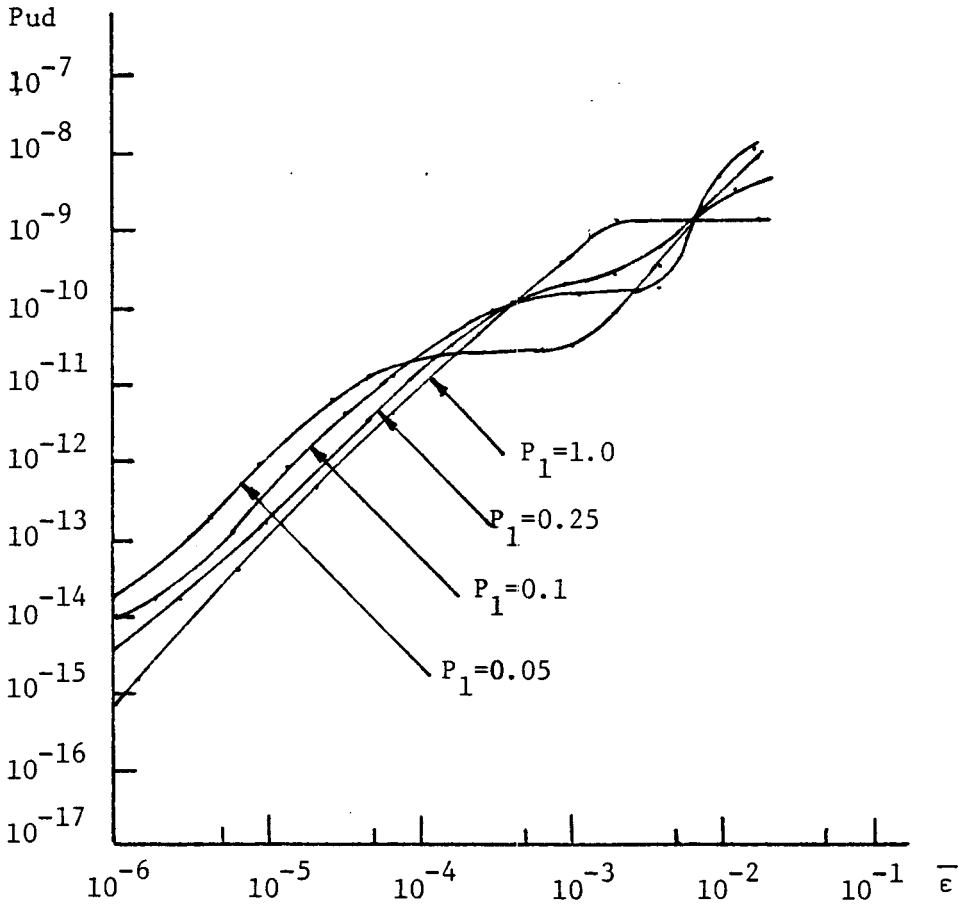


Fig. 5 Probability of undetected error of rate 1/2 selective-repeat ARQ with infinite receiver buffer.

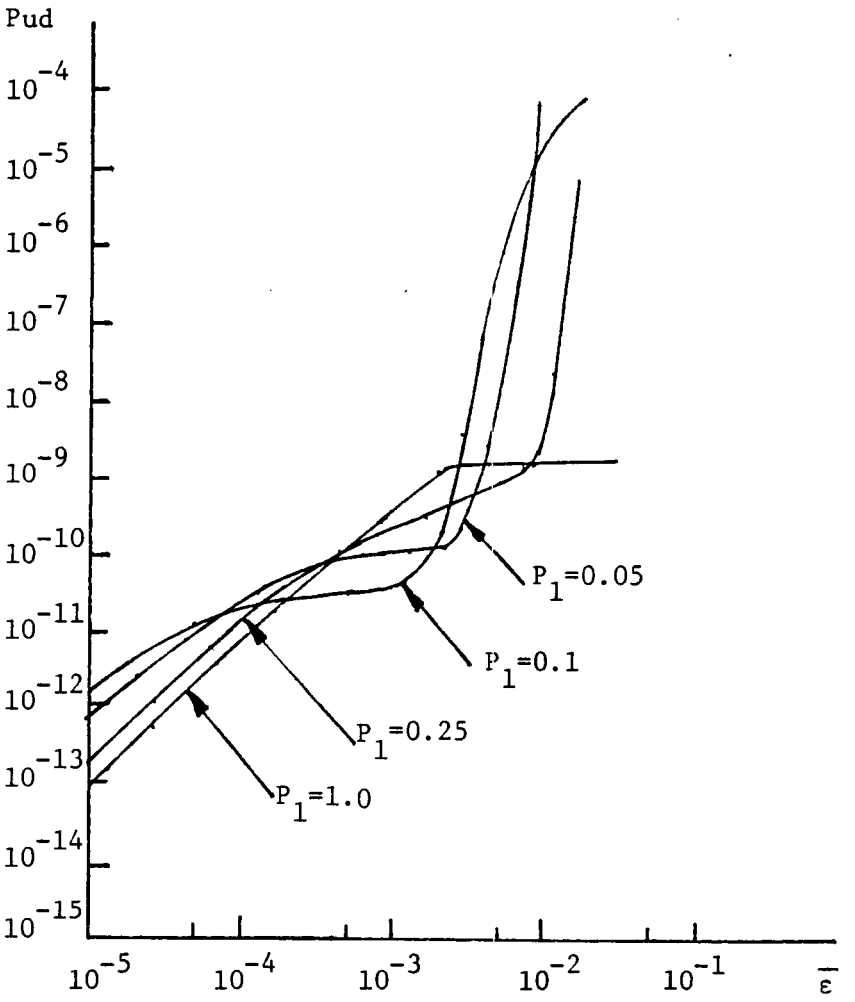


Fig. 6 Probability of undetected error of rate 1/2 selective-repeat ARQ with receiver buffer size B=5d.